

Towards the routine use of subdominant harmonics in gravitational-wave inference: re-analysis of GW190412 with generation X waveform models

Marta Colleoni,^{1,*} Maite Mateu-Lucena,¹ Héctor Estellés,¹ Cecilio García-Quirós,¹
David Keitel,¹ Geraint Pratten,² Antoni Ramos-Buades,^{3,1} and Sascha Husa¹

¹*Departament de Física, Universitat de les Illes Balears, IAC3 – IEEC, Crta. Valldemossa km 7.5, E-07122 Palma, Spain*

²*School of Physics and Astronomy and Institute for Gravitational Wave Astronomy,
University of Birmingham, Edgbaston, Birmingham, B15 9TT, United Kingdom*

³*Max Planck Institute for Gravitational Physics (Albert Einstein Institute), Am Mühlenberg 1, Potsdam, 14476, Germany*

(Dated: October 14, 2020)

We re-analyse the gravitational-wave event GW190412 with state-of-the-art phenomenological waveform models. This event, which has been associated with a black hole merger, is interesting due to the significant contribution from subdominant harmonics. We use both frequency-domain and time-domain waveform models. The PhenomX waveform models constitute the fourth generation of frequency-domain phenomenological waveforms for black hole binary coalescence; they have more recently been complemented by the time-domain PhenomT models, which open up new strategies to model precession and eccentricity, and to perform tests of general relativity with the phenomenological waveforms approach. Both PhenomX and PhenomT have been constructed with similar techniques and accuracy goals, and due to their computational efficiency this “generation X” model family allows the routine use of subdominant spherical harmonics in Bayesian inference. We show the good agreement between these and other state-of-the-art waveform models for GW190412, and discuss the improvements over the previous generation of phenomenological waveform models. We also discuss practical aspects of Bayesian inference such as run convergence, variations of sampling parameters, and computational cost.

PACS numbers: 04.30.-w, 04.80.Nn, 04.25.D-, 04.25.dg 04.25.Nx,

I. INTRODUCTION

The analysis of gravitational wave (GW) data from compact binary coalescences (CBCs) has long focused on the dominant quadrupole spherical harmonics, and the use of sub-dominant spherical harmonics has only recently started to play a prominent role for observational results [1–4] with the ground-based detectors Advanced LIGO [5] and Advanced Virgo [6]. The use of waveform models including sub-dominant harmonics can often break degeneracies between source parameters and improve the overall parameter estimation results even if the content of sub-dominant harmonics is weak for the most likely set of parameters, since the additional harmonics help to exclude parameter-space regions which are not consistent with the data under the more complete models.

Here we argue that models including subdominant harmonics should now be used routinely in GW parameter estimation, and how this is facilitated by the computational efficiency and accuracy of the “generation X” of phenomenological waveform models: the frequency-domain IMR_{PHENOMX} models [7–10], and the complementary time-domain IMR_{PHENOMT} models [11, 12]. In particular we demonstrate the capability of the generation X waveform family to deliver a suite of Bayesian inference results both quickly, and with relatively low total computational cost. One interesting application is on measuring the source distance, where breaking the degeneracy with the binary’s inclination through inclusion of subdominant harmonics

significantly improves accuracy [1, 13]. A quick turnaround on such precise distance measurements is particularly important in the context of electromagnetic counterparts, which have recently become more interesting also in the context of binary black hole (BBH) mergers, as discussed in connection with the discovery of GW190521 [3, 4]. Moreover, being able to efficiently compare posteriors for multiple models from the same family including different amount of physics (aligned spins only vs. precession, dominant modes only vs. inclusion of subdominant modes) allows detailed studies of waveform modelling systematics, increasing the confidence in the final parameter estimates for an event.

To provide a guide to the routine use of subdominant harmonics in practical parameter estimation and to demonstrate the capabilities of the “generation X” phenomenological waveform models on a specific example, in this paper we present improved parameter estimation results for the event GW190412 first published in [1]. GW190412 is special among the CBC observations reported by LIGO and Virgo to date [1–3, 14, 15] as the first BBH signal with clear evidence for significantly unequal component masses. Since the effect of sub-dominant spherical harmonic modes is stronger in unequal-mass systems, this event also provided the first observational evidence for their presence. On the other hand, as for previous BBH detections, only limited amounts of information about the spins of the black holes could be extracted in [1]. The event is of astrophysical interest as the unusual mass ratio hints at a more diverse population of merging BBHs in the Universe than the ‘vanilla’ events previously observed [14, 16] and starts to provide discriminating evidence between possible formation channels for

* marta.colleoni@uib.es

these systems [1, 17–25].

In general, the measurement of CBC source properties through matched filtering and Bayesian inference relies crucially on the quality of the waveform models used as the templates. Waveform models are typically synthesized from perturbative results (notably the post-Newtonian [26] and effective-one-body [27–29] frameworks and results for the ringdown frequencies of Kerr black holes [30]), and catalogues of numerical relativity (NR) simulations, such as the large catalogue of waveforms from the SXS collaboration [31, 32]. Several such models of varying complexity have been used in [1] to measure the source properties of GW190412. None of those models includes orbital eccentricity, but two of them include both precession and subdominant harmonics: the frequency-domain model IMRP_{PHENOM}PV3HM [33] and the time-domain model SEOBNRv4PHM [34]. Both are members of families of models which also include simpler models without precession or subdominant harmonics, which have been used for model comparison to determine the evidence for the presence of these effects in the data. Systematic errors in such waveform models, in particular regarding precession and sub-dominant harmonics, are not yet well understood, and the high computational cost for both the IMRP_{PHENOM}PV3HM and SEOBNRv4PHM models is one reason why investigations of potential systematics are challenging. None of the models used to analyze GW190412 in [1] is calibrated to precessing NR waveforms, and only in SEOBNRv4PHM have the subdominant harmonics been calibrated to NR, with IMRP_{PHENOM}PV3HM using approximate scalings to include their effects.

Our re-analysis of GW190412 focuses on the IMRP_{PHENOM}X family of inspiral-merger-ringdown waveform models [7–10] which constitute a thorough upgrade of previous versions of the family of frequency-domain phenomenological waveform models [33, 35–40] that is currently routinely used in CBC data analysis, including the IMRP_{PHENOM}PV3HM model mentioned above. In addition we study this event with the new IMRP_{PHENOM}T family of phenomenological time-domain waveforms (with the versions currently implemented in the LALSuite [41] package not yet including precession). An alternative re-analysis focusing on the NRSUR7DQ4 [42] model has been presented in [43]. NRSUR7DQ4 is based on reduced-order-modelling (ROM) [44] and has been calibrated to numerical relativity waveforms. It does however not span a sufficiently wide frequency range to cover the entire LIGO and Virgo bands for the mass range of GW190412, and [43] studies variations of the lower cutoff frequency.

None of the IMRP_{PHENOM}X models have yet been calibrated to precessing NR waveforms, but IMRP_{PHENOM}XHM, IMRP_{PHENOM}XPHM and IMRP_{PHENOM}THM are calibrated to subdominant harmonics from NR waveforms. The modularity and flexibility of the model family allows to compare different approximations for the effects of spin precession and different choices for the final spin of the remnant black hole. Furthermore the drastically reduced computational cost of the new waveforms allows us to test in detail the impact of varying some of the settings of the Bayesian sampling algorithms we

use [45–47]. We find that by replacing IMRP_{PHENOM}PV3HM with our upgraded precessing higher-modes model IMRP_{PHENOM}XPHM, the disagreement between the frequency and time-domain models observed in [1] can be reduced significantly, and the uncertainty intervals for key parameters can be tightened. The agreement between frequency and time-domain models is further improved when assuming that the source does not show significant spin-precession, which we confirm to be consistent with the observational data. A summary of our parameter estimation results and a comparison with the results from [1] can be found in Table I and is discussed below.

The paper is organized as follows. In Sec. II we collect preliminaries: remarks on notation, a summary of the results found in [1] on the event GW190412 and brief descriptions of the different waveform models and Bayesian inference methods used there and in the present paper. We then present our main parameter estimation results on GW190412 in Sec. III, further investigations on systematic and sampling errors in Sec. IV, and a summary and our conclusions in Sec. V. Further comparisons are presented in appendices: while all our main results are obtained with the parallel Bilby code [45–47], in appendix A we also compare these results with comparison runs of the LALInference code [48]. And in appendix B we compare our main Bilby runs, which have been obtained with marginalisation over distance in order to improve convergence, with runs that do not use this approximation. Different IMRP_{PHENOM}X implementations of precession and approximations for the spin of the merger remnant are studied in appendix C. Finally, we compare results in more detail against the IMRP_{PHENOM}PV3HM model in appendix D and study the impact of alternative spin priors in appendix E.

Posterior samples from our preferred run for each waveform model are released in [49].

II. PRELIMINARIES

A. Notation and conventions

We will report all masses in units of the solar mass M_{\odot} . Masses are reported both in the detector frame, where they appear redshifted, and in the source frame, assuming a standard cosmology [50] (see Appendix B of [14]). We will report source-frame masses with a superscript s , as in m_1^s to denote the mass of the larger black hole in the source frame. We will drop the superscript to denote masses in the detector frame, and to represent general relations between different mass parameters. Individual component masses are denoted by m_i , the total mass is $M = m_1 + m_2$, and the chirp mass by $\mathcal{M} = (m_1 m_2)^{3/5} M^{-1/5}$. The mass ratio is defined as $q = m_2/m_1 \leq 1$.

We also define two effective spin parameters which are commonly used in waveform modelling and parameter estimation. First, the parameter χ_{eff} [27, 51, 52] is defined as

$$\chi_{\text{eff}} = \frac{m_1 \chi_1 + m_2 \chi_2}{m_1 + m_2}, \quad (2.1)$$

where the χ_i are the projections of the spin vectors of the individual black holes onto the orbital angular momentum. Second, the effective spin precession parameter χ_p [53] has been designed to capture the dominant effect of precession. It corresponds to an approximate average over many precession cycles of the spin in the precessing orbital plane, and is defined in terms of the average spin magnitude S_p , [37, 53]

$$S_p = \frac{1}{2} (A_1 S_{1,\perp} + A_2 S_{2,\perp} + |A_1 S_{1,\perp} - A_2 S_{2,\perp}|), \quad (2.2)$$

$$= \max(A_1 S_{1,\perp}, A_2 S_{2,\perp}), \quad (2.3)$$

where $A_1 = 2 + 3/(2q)$, and χ_p is then defined as

$$\chi_p = \frac{S_p}{A_1 m_1^2}. \quad (2.4)$$

Both χ_{eff} and χ_p are dimensionless and thus independent of the frame (source or detector).

Throughout this work we will employ waveforms with several multipoles beyond the quadrupolar contribution. Unless otherwise stated, we will consider pairs of both positive and negative modes when referring to a particular multipole. For example, to refer to a set of multipoles $(l, m) = (2, \pm 2), (2, \pm 1)$ we will use the simplified notation $(l, |m|) = (2, 2), (2, 1)$ or simply $(2, 2), (2, 1)$.

B. Summary of the event GW190412

GW190412 is the first BBH detection from the O3 observing run published [1] by the LIGO-Virgo Collaboration (LVC). It was observed by all three detectors, with a network signal-to-noise ratio (SNR) of $19.1_{-0.3}^{+0.1}$ from the final coherent Bayesian analysis reported in [1]. In briefly summarizing the event properties, we concentrate here on the results from Bayesian inference using the Bilby [45–47], LALInference [48] and RIFT [54, 55] packages and the SEOBNRv4PHM [34] and IMRPHENOMPv3HM [33] waveform models. These LVC estimates of the properties of GW190412’s source are listed in Table I together with our own results.

In summary, GW190412 came from a BBH with individually unremarkable source-frame masses $m_1^s \approx 30M_\odot$ and $m_2^s \approx 8M_\odot$, but the mass ratio $q = m_2/m_1 \approx 0.3$ (< 0.5 at 99% probability) is much lower than inferred for any previous detection. The system’s effective spin parameter appears to be low, but significantly different from zero: $\chi_{\text{eff}} = 0.25_{-0.11}^{+0.08}$ when combining results from both waveforms. The effective precession parameter has also been constrained, but more weakly, to $\chi_p = 0.31_{-0.16}^{+0.19}$. While there is some information gain compared to the prior on this quantity, and the posterior is peaked away from zero, it prefers lower values than the prior. The results from SEOBNRv4PHM and IMRPHENOMPv3HM agree within 90% uncertainties for all quantities reported in [1], but do show some qualitative differences – for example,

the IMRPHENOMPv3HM posteriors tend to less unequal q together with lower χ_{eff} .

Additional inference runs were also reported in [1] for several waveform models with reduced physics content (see their Table I and references therein), with the goal of estimating the evidence for the presence of precession and HMs in the observed signal. Clear and robust evidence was found for the presence of HMs, with \log_{10} Bayes factors between 3.0 and 4.1 in favour of models including HMs over those only including $\ell = 2$ modes, depending on the waveform model family, sampling method and whether precession was also included at the same time. On the other hand, there was no clear evidence for or against precession, with the obtained Bayes factors for that hypothesis test remaining within systematic uncertainties for each waveform model family. We summarise Bayes factors listed in the LVC publication, together with those obtained from our own analysis, in Table IV.

The main astrophysical conclusions that [1] drew from GW190412 include that the event’s properties, especially its mass ratio, are somewhat unexpected for draws from a BBH population as inferred from the previous two observing runs [14, 16], but not in clear tension with it. Furthermore, the formation of the source system challenges some astrophysical models that mostly predict mergers with near-equal masses, but GW190412 is still compatible with most versions of both isolated binary evolution and dynamical assembly [56, 57]. Implications for specific formation scenarios have since been the topic of many studies [17–24], and more accurate and robust inference of the system’s mass and spin parameters can be crucial in further constraining these channels.

Additional PE analyses of GW190412 have also been reported by [58], focusing on the impact of alternative spin priors, and recently by [43], focusing on the NRSUR7DQ4 model and the effect of varying the lower cutoff frequency.

C. Waveform models used

CBC parameter estimation currently mostly uses waveform models from three different families:

- Models constructed within the effective-one-body (EOB) framework [27–29, 62]. In a first stage these are constructed as time-domain models, where Hamiltonians and GW fluxes are calibrated to NR simulations, and the ordinary differential equations resulting from the Hamiltonian equations are solved numerically for the inspiral, and carried through the merger and ringdown with phenomenological models. One of two models used in [1] and describing both precession and subdominant harmonics, SEOBNRv4PHM [34], and its restriction to the dominant quadrupole content, SEOBNRv4P, belong to this family. These models are typically computationally expensive, thus it is common to produce reduced-order-models (ROMs) to accelerate the evaluation of the waveforms. Two such models have been used in [1]: SEOBNRv4_ROM, which describes non-precessing systems including HMs, and SEOBNRv4HM_ROM, which

TABLE I. Inferred parameter values for GW190412 and their 90% credible intervals, obtained using precessing models including higher multipoles. Columns 2–4 correspond to the results from the LVC analyses [1], the fifth column gives the new results from our preferred inference run, which is run 26 from Table III with the precessing higher-modes model IMRP_{PHENOM}XPHM, and the last column combines the LVC SEOBNRv4PHM results with IMRP_{PHENOM}XPHM in the same way that they were combined with IMRP_{PHENOM}Pv3HM in [1].

parameter	SEOBNRv4PHM	IMRP _{PHENOM} Pv3HM	LVC Combined	IMRP _{PHENOM} XPHM	Combined
m_1^s/M_\odot	$31.7^{+3.6}_{-3.5}$	$28.1^{+4.8}_{-4.3}$	$29.7^{+5.0}_{-5.3}$	$30.0^{+5.2}_{-4.3}$	$30.9^{+4.3}_{-4.4}$
m_2^s/M_\odot	$8.0^{+0.9}_{-0.7}$	$8.8^{+1.5}_{-1.1}$	$8.4^{+1.8}_{-1.0}$	$8.4^{+1.3}_{-1.1}$	$8.2^{+1.2}_{-0.9}$
M^s/M_\odot	$39.7^{+3.0}_{-2.7}$	$36.9^{+3.7}_{-2.9}$	$38.1^{+4.0}_{-3.7}$	$38.4^{+4.2}_{-3.7}$	$39.1^{+3.6}_{-3.3}$
\mathcal{M}^s/M_\odot	$13.3^{+0.3}_{-0.3}$	$13.2^{+0.5}_{-0.3}$	$13.3^{+0.4}_{-0.3}$	$13.3^{+0.5}_{-0.4}$	$13.3^{+0.4}_{-0.3}$
q	$0.25^{+0.06}_{-0.04}$	$0.31^{+0.12}_{-0.07}$	$0.28^{+0.13}_{-0.06}$	$0.28^{+0.09}_{-0.07}$	$0.27^{+0.09}_{-0.06}$
M_f/M_\odot	$38.6^{+3.1}_{-2.8}$	$35.7^{+3.8}_{-3.0}$	$37.0^{+4.1}_{-3.9}$	$37.2^{+4.3}_{-3.3}$	$38.0^{+3.5}_{-3.3}$
χ_f	$0.68^{+0.04}_{-0.04}$	$0.67^{+0.07}_{-0.07}$	$0.67^{+0.05}_{-0.05}$	$0.67^{+0.05}_{-0.05}$	$0.67^{+0.05}_{-0.04}$
m_1/M_\odot	$36.5^{+4.2}_{-4.2}$	$32.3^{+5.7}_{-5.2}$	$34.2^{+5.7}_{-6.5}$	$34.4^{+6.2}_{-5.1}$	$35.5^{+5.2}_{-5.1}$
m_2/M_\odot	$9.2^{+0.9}_{-0.7}$	$10.1^{+1.6}_{-1.2}$	$9.7^{+1.8}_{-1.1}$	$9.6^{+1.4}_{-1.2}$	$9.4^{+1.3}_{-0.9}$
M/M_\odot	$45.7^{+3.5}_{-3.3}$	$42.5^{+4.4}_{-3.7}$	$43.9^{+4.6}_{-4.7}$	$44.1^{+5.0}_{-3.7}$	$44.9^{+4.3}_{-3.8}$
\mathcal{M}/M_\odot	$15.3^{+0.1}_{-0.2}$	$15.2^{+0.3}_{-0.2}$	$15.3^{+0.2}_{-0.2}$	$15.3^{+0.4}_{-0.2}$	$15.3^{+0.3}_{-0.2}$
χ_{eff}	$0.28^{+0.08}_{-0.08}$	$0.22^{+0.08}_{-0.11}$	$0.25^{+0.08}_{-0.11}$	$0.25^{+0.1}_{-0.1}$	$0.26^{+0.08}_{-0.09}$
χ_p	$0.31^{+0.14}_{-0.15}$	$0.31^{+0.24}_{-0.17}$	$0.30^{+0.19}_{-0.15}$	$0.23^{+0.20}_{-0.13}$	$0.27^{+0.17}_{-0.15}$
χ_1	$0.46^{+0.12}_{-0.15}$	$0.41^{+0.22}_{-0.24}$	$0.43^{+0.16}_{-0.26}$	$0.39^{+0.16}_{-0.17}$	$0.43^{+0.14}_{-0.17}$
D_L/Mpc	740^{+120}_{-130}	740^{+150}_{-190}	730^{+140}_{-170}	734^{+161}_{-187}	737^{+141}_{-159}
z	$0.15^{+0.02}_{-0.02}$	$0.15^{+0.03}_{-0.04}$	$0.15^{+0.03}_{-0.03}$	$0.15^{+0.03}_{-0.04}$	$0.15^{+0.03}_{-0.03}$
$\hat{\theta}_{JN}$	$0.71^{+0.23}_{-0.21}$	$0.71^{+0.39}_{-0.27}$	$0.73^{+0.34}_{-0.24}$	$0.75^{+0.36}_{-0.28}$	$0.73^{+0.31}_{-0.24}$
ρ_H	$9.5^{+0.1}_{-0.2}$	$9.5^{+0.2}_{-0.3}$	$9.5^{+0.1}_{-0.3}$	$9.4^{+0.2}_{-0.3}$	$9.4^{+0.2}_{-0.3}$
ρ_L	$16.2^{+0.1}_{-0.2}$	$16.1^{+0.2}_{-0.3}$	$16.2^{+0.1}_{-0.3}$	$16.1^{+0.2}_{-0.3}$	$16.2^{+0.2}_{-0.3}$
ρ_V	$3.7^{+0.2}_{-0.5}$	$3.6^{+0.3}_{-1.0}$	$3.6^{+0.3}_{-1.0}$	$3.6^{+0.3}_{-0.8}$	$3.7^{+0.3}_{-0.7}$
ρ_{HLV}	$19.1^{+0.2}_{-0.2}$	$19.0^{+0.2}_{-0.3}$	$19.1^{+0.1}_{-0.3}$	$18.9^{+0.2}_{-0.3}$	$19.0^{+0.2}_{-0.4}$

corresponds to the $l = |m| = 2$ content. Several generations of these models have been built, and [1] uses the fourth generation (“v4”).

- Phenomenological models, which are constructed as piecewise closed-form expressions that are calibrated to post-Newtonian or EOB inspiral descriptions and NR waveforms, and which can be evaluated very rapidly. As for the EOB models, several generations of such models have been built, with the generation used in [1] all constructed from the baseline IMRP_{PHENOM}D model for the $l = |m| = 2$ modes of non-precessing binaries. Analytical approximate maps are used to model the HM content and precession. We will refer to the phenomenological models used in [1] as the *third generation* (counting IMRP_{PHENOM}A [64] as the first generation, and IMRP_{PHENOM}B [51], IMRP_{PHENOM}C [52] and IMRP_{PHENOM}P [37] as the second generation). The IMRP_{PHENOM}X family constitutes the next generation and current state of the art for such models, and corresponds to an update of essentially all aspects of the model. We will provide further details on how it relates to the third generation below. While previous phenomenological waveform models have all been constructed in the frequency domain, IMRP_{PHENOM}X is complemented by the new IMRP_{PHENOM}T time-domain models [11, 12]. Currently, only non-precessing versions of IMRP_{PHENOM}T have been implemented in LALSuite

[41] and can be used for the analysis presented here (see however [11] for a Mathematica implementation lacking sub-dominant harmonics).

- Finally, ROMs have also been successfully applied directly to interpolate between NR or hybrid waveforms in the time domain. Hybrids are built from appropriately “gluing” NR waveforms to an early inspiral described by an EOB model. The latest models of this kind are the non-precessing NRHYBSUR3DQ8 [65], which has been used in [1], and the precessing NRSUR7DQ4 [42], which has not been used in that original discovery paper, since it does not span a sufficiently wide frequency range to cover the entire LIGO and Virgo bands for the mass range of GW190412. See however [43] for results obtained more recently with NRSUR7DQ4, and studies on varying the lower cutoff frequency.

For a complete list of all the waveform models used in [1] and the present paper see Table II. We now turn to describing the new IMRP_{PHENOM}X and IMRP_{PHENOM}T families in more detail.

The starting point of the IMRP_{PHENOM}X family is IMRP_{PHENOM}XAS [7], which models the $(\ell, |m|) = (2, 2)$ modes of signals from non-precessing BBHs, and has been extended to include the $(\ell, |m|) = (2, 1), (3, 3), (3, 2), (4, 4)$ modes by IMRP_{PHENOM}XHM. All these modes have been calibrated to

TABLE II. Waveform models used in this paper. We indicate which models include precession and which multipoles are included for each model. For precessing models, the multipoles correspond to those in the co-precessing frame.

family	full name	precession	multipoles ($\ell, m $)	ref.
EOBNR	SEOBNRv4HM_ROM	×	(2, 2)	[59]
	SEOBNRv4_ROM	×	(2, 2), (2,1), (3, 3), (4, 4), (5,5)	[29, 60]
	SEOBNRv4P	✓	(2, 2), (2, 1)	[34, 61, 62]
	SEOBNRv4PHM	✓	(2, 2), (2, 1), (3, 3), (4, 4), (5,5)	[34, 61, 62]
Phenom - Gen. 3	IMRPHENOMD	×	(2, 2)	[35, 36]
	IMRPHENOMHM	×	(2, 2), (2, 1), (3, 3), (3, 2), (4,4), (4, 3)	[39]
	IMRPHENOMPv2	✓	(2, 2)	[37, 38]
	IMRPhenomPv3	✓	(2, 2)	[40]
	IMRPHENOMPv3HM	✓	(2, 2), (2, 1), (3, 3), (3, 2), (4, 4), (4, 3)	[33]
NR surrogate	NRHYBSUR3DQ8	×	$\ell \leq 4, (5, 5)$ but not (4, 0), (4, 1)	[63]
PhenomX	IMRPHENOMXAS	×	(2, 2)	[7]
	IMRPHENOMXHM	×	(2, 2), (2, 1), (3, 3), (3, 2), (4,4)	[8, 9]
	IMRPHENOMXP	✓	(2, 2)	[10]
	IMRPHENOMXPHM	✓	(2, 2), (2, 1), (3, 3), (3, 2), (4, 4)	[10]
PhenomT	IMRPHENOMT	×	(2, 2)	[11]
	IMRPHENOMTHM	×	(2, 2), (2, 1), (3, 3), (4,4), (5,5)	[12]

around 500 numerical waveforms, as reported in [9], whereas for the third generation of Phenom waveforms only the dominant $(\ell, |m|) = (2, 2)$ modes had been calibrated to 19 NR waveforms in 2015 [35, 36]. IMRPHENOMXPHM [10] is our most complete state-of-the-art phenomenological waveform model for quasi-circular precessing signals including the same modes as IMRPHENOMXHM, and IMRPHENOMXP is the corresponding precessing version of the dominant-mode IMRPHENOMXAS.

Following the paradigm [37, 66, 67] adopted by previous phenomenological models, the precessing versions are obtained by extending the non-precessing waveforms by an approximate map, which identifies the non-precessing spherical harmonic modes with the precessing modes in a non-inertial co-precessing frame. We refer to this procedure as “twisting up”. For a detailed discussion of the procedure and the conventions employed to describe precessing waveforms in the frequency domain see [10]. The twisting construction relies on a prescription for the three frequency-dependent Euler angles which rotate the modes in the non-inertial frame into the inertial frame, which is used for gravitational wave data analysis. These angles encode the amplitude and phase modulations determined by the precession dynamics and are by default computed following the double spin multiple-scale-analysis (MSA) prescription [68], although the model has the option to instead evolve them at next-to-next-to-leading order (NNLO) in the post-Newtonian approximation as a function of an effective single spin. The MSA prescription is similar to that used in IMRPHENOMPv3, and the NNLO prescription to what is used in IMRPHENOMPv2. The mode content of IMRPHENOMXPHM in the co-precessing frame can be freely specified, with the full set comprising the modes $(\ell, |m|) = (2, 2), (2, 1), (3, 3), (3, 2), (4, 4)$. Note also that the modelling of the co-precessing $(3, |2|)$ mode incorporates

mode-mixing effects. Another notable feature of IMRPHENOMXPHM is its use of the multibanding interpolation method introduced by [69], implemented in an improved way as described in [9]. The algorithm is applied to the evaluation of both the aligned-spin waveforms and the precession angles, with default thresholds that can be further relaxed by the user to allow for even quicker exploratory runs. In Sec. III D, we will discuss the benefits of multibanding for parameter estimation studies.

The “twisting-up” procedures adopted by IMRPHENOMPv3HM and IMRPHENOMXPHM are almost equivalent, as both models adopt by default the MSA prescription. However, IMRPHENOMXPHM by default falls back to NNLO angles when MSA failures are encountered. This feature is advantageous in parameter estimation studies as, with no fallback in place, samplers may get stuck in regions of parameter space where the MSA equations become numerically unstable. IMRPHENOMXPHM offers several options to set the precession prescription, with the default version (also called version 223) being the one just described. Another relevant version for this work will be version 102, which instead enforces the use of NNLO angles. IMRPHENOMX also allows to choose among four final-spin formulas, with the default version using a precession-averaged equation inspired by the MSA formalism (“FS version 3”), see [10] for details. Alternative versions attach the in-plane spins to the larger mass, either relying on the usual effective precession spin χ_p (version 0, which is adopted by all third generation Phenom models), or by taking the norm of the in-plane spin vectors at the reference frequency (version 2). We will discuss the impact of these settings on parameter estimation in appendix C.

A further important element distinguishing IMRPHENOMXPHM from its predecessor IMRPHENOMPv3HM is the calibra-

tion of the underlying aligned-spin higher modes (HM) to NR waveforms. It has been already shown in [9] that this results in an increased faithfulness both with respect to NR and to NRSUR7DQ4. We will study the impact of this HM calibration on parameter estimation in appendix D.

Thanks to the modular way that IMRP_{PHENOM}XHM and IMRP_{PHENOM}XPHM are constructed from the baseline dominant-mode model IMRP_{PHENOM}XAS, their LALSuite implementations when called with the dominant (2, 2) modes only will reproduce the IMRP_{PHENOM}XAS and IMRP_{PHENOM}XP models respectively. This approach has the added advantage of providing multibanding speedup, and we used it for most of the dominant-mode PE runs included in this paper. For clarity those runs are usually referred to as “IMRP_{PHENOM}X(P)HM (2,2)” (or shortened versions of this) in the following.

Recently, IMRP_{PHENOM}X models have been complemented by the new IMRP_{PHENOM}T time-domain models [11, 12]. IMRP_{PHENOM}T and IMRP_{PHENOM}THM model the $(\ell, |m|) = (2, 2), (2, 1), (3, 3), (4, 4), (5, 5)$ modes of non-precessing signals, while IMRP_{PHENOM}TP describes the $\ell = 2$ sector of precessing signals applying the “twisting-up” procedures to the dominant mode $(\ell, |m|) = (2, 2)$ described by IMRP_{PHENOM}T. Like the corresponding non-precessing models of the IMRP_{PHENOM}X family, IMRP_{PHENOM}T and IMRP_{PHENOM}THM have been calibrated to around 500 numerical waveforms. Model construction follows a similar design as in other phenomenological models, describing the inspiral region of the signal through a post-Newtonian quasi-circular approximant, TaylorT3 [70], extended to higher pseudo-PN orders calibrated with numerical waveforms, providing phenomenological descriptions of the merger signal and including a calibrated ringdown description based on the quasinormal mode expansion of the signal [71].

D. Methodology for our parameter estimation analysis

For this reanalysis, we use v2 of the strain data [72] for GW190412 released through the Gravitational Wave Open Science Center (GWOSC) [73, 74], with a default sampling rate of 16384 Hz, for consistency with the official LVC study. This version has non-linear subtraction [75] of 60 Hz power lines applied to it. We also use the power spectral densities (PSDs) [76, 77] and calibration uncertainties [78] included in v11 of the posterior sample release [79] for the event. We analyse 8 s of strain data from each of the Hanford, Livingston and Virgo detectors around the trigger time of the event, as reported in GraceDB [80].

We have carried out most of our Bayesian parameter estimation runs using the python-based package pBilby [45, 81, 82] with static nested sampling [83] as implemented in the dynesty Python code [84]. We use the default parameters of the dynesty implementation in Bilby, apart from the following more important parameters: We vary the number of nested sampling live points (n_{live}) and the minimum length of the chain, in terms of multiples (n_{act}) of its auto-correlation length. We also set a minimal ($walks$) and maximal ($maxmcmc$) number of Markov-

Chain Monte Carlo (MCMC) steps. All parallel Bilby runs are carried out with four statistically independent sampling runs, usually referred to as “seeds”.

The lower and upper cutoff frequencies for the likelihood integration were taken to be 20 Hz and 2048 Hz respectively. We adopted the same “tight” priors as employed for the LVC pBilby runs, with narrow bounds based on initial exploratory runs. Explicit prior settings are listed in appendix E. We have also examined the impact of alternative spin priors, as explained in appendix E. Besides custom post-processing scripts, the PESummary [85] package was also used for comparisons of multiple runs.

In order to claim confidence in our results, we need to check how posteriors change when refining our sampler settings. To this end we have varied the number of live points n_{live} and the number n_{act} of autocorrelation times to use before accepting a point. Our main series of runs, summarized in Table III, uses marginalisation over the distance parameter and then a reconstruction of the distance posterior as discussed in [47, 86] to accelerate the convergence of nested sampling¹. In Sec. B we compare the distance-marginalised results with a series of simulations that do not use distance marginalisation (listed in Table VIII).

Some additional comparison runs as listed in Table VII were performed with the C library LALInference (LI) [48], using both MCMC and nested sampling, since the parallel Bilby code is still relatively new. The LALInference runs do not use distance (or other types of) marginalisation and they are compared with our default series of runs in appendix A.

To quantify the differences between results which use different sampler settings, we also compute their Jensen-Shannon (JS) divergence [87] and analyze the results in Sec. III D. This quantity measures the distance between two probability distributions as a value in [0,1], where 0 means that both distributions are exactly alike and 1 means maximal divergence. The JS divergence between two distributions $p(x)$ and $q(x)$ is defined as

$$D_{\text{JS}}(p|q) = \frac{1}{2} \left[D_{\text{KL}} \left(p \left| \frac{1}{2}(p+q) \right. \right) + D_{\text{KL}} \left(\frac{1}{2}(p+q) \left| q \right. \right) \right], \quad (2.5)$$

where the Kullback-Leibler divergence D_{KL} [88] is defined as

$$D_{\text{KL}}(p|q) = \int p(x) \log_2 \left(\frac{p(x)}{q(x)} \right) dx. \quad (2.6)$$

III. PE RESULTS FOR GW190412

A. Analysis with IMRP_{PHENOM}XPHM

We first compare posterior distributions for the two precessing models used in the LVC paper [1], SEOBNRv4PHM and

¹ We do not use time-marginalization, due to potential issues in the sky-localization posteriors. We also do not use phase-marginalization since it is not fully compatible with precessing and higher-modes waveforms.

Approximant	Run n	Modes (l, $ m $)	PV FS PMB MB	Prior	n_{live}	n_{act}	CPU h	L. eval.	Cost/L. eval. [ms]
IMRP _{HENOMXHM}	1	(2,2)	- - - D	Aligned spin	2048	10	1186	1.10×10^8	38.96
	2	(2,2)	- - - D	Aligned spin	2048	50	5819	5.48×10^8	38.26
	3	(2,2)	- - - 10^{-1}	Aligned spin	2048	10	1171	1.09×10^8	38.83
	4	(2,2)	- - - 10^{-2}	Aligned spin	2048	10	1166	1.09×10^8	38.43
	5	D	- - - D	Aligned spin	2048	10	1666	1.29×10^8	46.39
	6	D	- - - D	Aligned spin	2048	50	7918	6.38×10^8	44.71
	7	D	- - - 10^{-1}	Aligned spin	2048	10	1501	1.29×10^8	41.78
	8	D	- - - 10^{-2}	Aligned spin	2048	10	1494	1.28×10^8	42.15
IMRP _{HENOMTHM}	9	(2,2)	- - - -	Aligned spin	1024	10	874	5.06×10^7	62.22
	10	(2,2)	- - - -	Aligned spin	2048	10	1649	1.02×10^8	58.00
	11	D	- - - -	Aligned spin	1024	10	1226	6.19×10^7	71.27
	12	D	- - - -	Aligned spin	2048	10	2332	1.22×10^8	68.69
IMRP _{HENOMXPHM}	13	(2,2)	D D D D	Precessing	2048	10	1375	1.11×10^8	44.30
	14	(2,2)	D D D D	Precessing	2048	50	6681	5.54×10^8	43.42
	15	(2,2)	D D 10^{-1} 10^{-1}	Precessing	2048	10	1319	1.11×10^8	42.40
	16	(2,2)	D D 10^{-1} 10^{-2}	Precessing	2048	10	1320	1.13×10^8	41.89
	17	(2,2)	D D 10^{-2} 10^{-1}	Precessing	2048	10	1297	1.11×10^8	42.00
	18	(2,2)	D D 10^{-2} 10^{-2}	Precessing	2048	10	1296	1.31×10^8	41.54
	19	D	D D D D	Precessing	512	10	679	3.02×10^7	80.84
	20	D	D D D D	Precessing	512	50	3002	1.36×10^8	79.47
	21	D	D D D D	Precessing	1024	10	1318	6.11×10^7	77.62
	22	D	D D D D	Precessing	1024	50	6129	3.01×10^8	73.20
	23	D	D D D D	Precessing	2048	10	2670	1.32×10^8	72.95
	24	(2,2),(2,1)	D D D D	Precessing	2048	10	1580	1.19×10^8	47.77
	25	(2,2),(2,1),(3,2),(3,3)	D D D D	Precessing	2048	10	2325	1.04×10^8	80.44
	26	D	D D D D	Precessing	2048	50	13530	6.61×10^8	73.64
	27	D	D D D D	Precessing	4096	10	5718	2.79×10^8	73.79
	28	D	D D 10^{-1} 10^{-1}	Precessing	2048	10	1900	1.31×10^8	52.01
	29	D	D D 10^{-1} 10^{-2}	Precessing	2048	10	1984	1.33×10^8	53.70
	30	D	D D 10^{-2} 10^{-1}	Precessing	2048	10	1989	1.32×10^8	54.19
	31	D	D D 10^{-2} 10^{-2}	Precessing	2048	10	2001	1.32×10^8	54.63
	32	D	223 0 D D	Precessing	2048	10	2679	1.31×10^8	73.80
	33	D	223 1 D D	Precessing	2048	10	2711	1.03×10^8	94.83
	34	D	223 2 D D	Precessing	2048	10	2647	1.31×10^8	72.52
	35	D	102 0 D D	Precessing	2048	10	2217	1.29×10^8	61.49
	36	D	102 1 D D	Precessing	2048	10	2160	1.28×10^8	60.79
	37	D	102 2 D D	Precessing	2048	10	2230	1.30×10^8	61.68

TABLE III. This table lists all the runs we have performed with parallel Bilby and distance marginalisation on open data for GW190412. For each run, we indicate the LALSuite waveform approximant called along with various waveform settings: the mode content, precessing prescription (PV), final spin version (FS), and multibanding thresholds applied to the evaluation of the Euler angles (PMB) and of the aligned-spin modes (MB); as well as the prior used and the chosen sampler settings (n_{live} and n_{act}). We also provide the computational cost of each run, the number of likelihood evaluations and the mean cost of each evaluation. Two runs are highlighted in boldface: run 26 is our preferred IMRP_{HENOMXPHM} run for astrophysical results, while the cheaper but essentially equally accurate run 23 serves as a comparison baseline for alternative model choices and sampler settings.

IMRP_{HENOMPV3HM}, with our IMRP_{HENOMXPHM} model in Fig. 1. The posterior results for SEOBNRv4PHM and IMRP_{HENOMPV3HM} are taken from the official LVC release samples, while those shown for IMRP_{HENOMXPHM} correspond to our preferred run (run 26, shown in bold, in Table III), which uses $n_{\text{live}} = 2048$ and $n_{\text{act}} = 50$.

The parameters recovered with the three models considered here are broadly consistent, as can be seen by inspecting the joint posterior distributions for some of the key source properties (Fig. 1) as well as the corresponding maximum-likelihood waveforms (Fig. 2).

While all these models agree fairly well on the estimated distance and inclination of the source, differences are clearly visible in the mass and spin parameters. The component masses estimated with IMRP_{HENOMXPHM} lie in between those estimated with SEOBNRv4PHM and IMRP_{HENOMPV3HM}. We notice that IMRP_{HENOMPV3HM} tends to return broader posteriors than the other models, which appear to be somewhat more consistent with each other. We also see that SEOBNRv4PHM is favouring more asymmetric masses and higher χ_{eff} and χ_p , while IMRP_{HENOMXPHM} prefers lower values of χ_p , consistently with our analysis of Bayes factors, which will be pre-

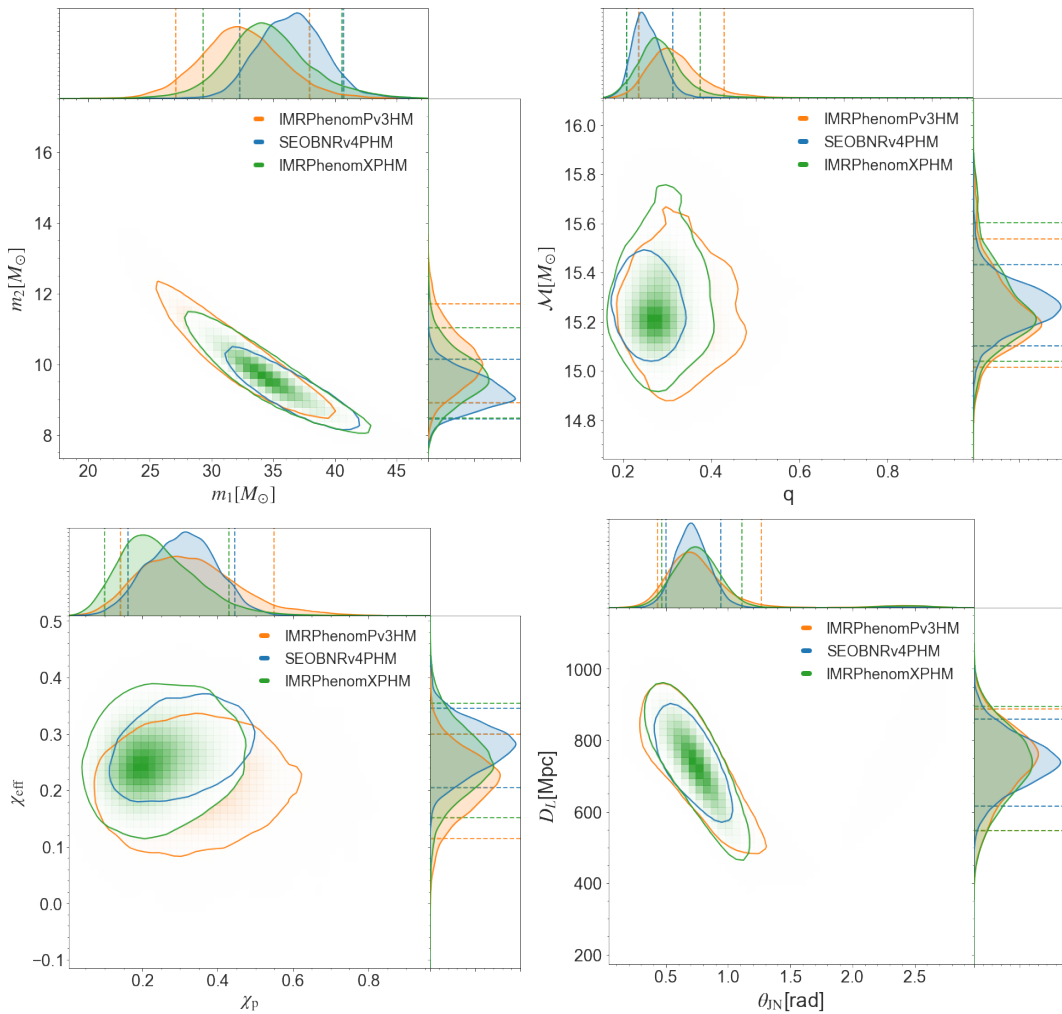


FIG. 1. Posterior distributions for the masses, effective spin parameters, distance and inclination of GW190412, as estimated with SEOBNRv4PHM (blue), IMRPHENOMPv3HM (orange) and IMRPHENOMXPHM (green) (using run 26 from Table III). For each pair of parameters, the square central panel shows the 2D joint posteriors with contours marking 90% credible regions; while the top and right side-panels show one-dimensional distributions for the individual parameters, with dashed lines indicating 90% credible intervals.

sented in the next section.

Given the better agreement between the SEOBNRv4PHM and IMRPHENOMXPHM models, we compute their combined posterior as our best estimate of the source parameters, and list the estimated source parameters and error estimates in Table I. These improvements of the LVC estimates which combine SEOBNRv4PHM and IMRPHENOMPv3HM posteriors constitute our main astrophysical result.

B. Comparison with models of reduced content

To investigate support in the GW190412 data for the presence of subdominant modes and precession, we compare a reference run using the full IMRPHENOMXPHM model (23 in Table III) against additional runs with equal sampler settings but using the models IMRPHENOMT, IMRPHENOMXAS, IMRPHENOMXP

which drop either one or both of these additional aspects. As discussed above, in fact we called IMRPHENOMXHM and IMRPHENOMXPHM with (2, 2)-modes only instead of the named XAS / XP LALSuite approximants.

Our main tool to compare how well each model fits the data are Bayes factors, i.e. ratios of the marginal likelihoods of the models to be compared. If we indicate by \mathcal{Z}_A the marginal likelihood of model A and by \mathcal{Z}_B that of model B, the Bayes factor will take the form

$$\mathcal{B}_{B/A} = \frac{\mathcal{Z}_B}{\mathcal{Z}_A}, \quad (3.1)$$

with $\log_{10} \mathcal{B}_{B/A} > 1$ indicating that model B is strongly preferred over model A. In Table IV we list $\log_{10} \mathcal{B}_{B/A}$ for two sets of hypotheses, namely whether the signal is best described 1) by an aligned-spin or precessing model and 2) by a quadrupole-only or higher-mode model. Error bars are computed using

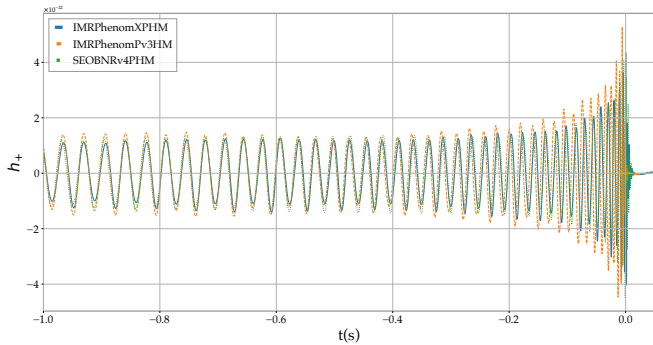


FIG. 2. Waveforms (h_+ polarization as a function of time) corresponding to the maximum-likelihood samples for IMRPHENOMXPHM (blue), IMRPHENOMPv3HM (orange) and SEOBNRv4PHM (green). The models show good agreement during the inspiral, while larger differences can be appreciated around merger. The maximum-likelihood sample for IMRPHENOMXPHM is taken from run 26.

error propagation, from the uncertainties associated to the evidence integral for each run.

We find strong evidence ($\log_{10} \mathcal{B} > 3$) for the presence of higher modes, consistent with the results from [1]. On the other hand, there is no actual evidence from the data to infer the presence of precession effects over aligned-only spins, neither when limiting to dominant multipoles or when including the subdominant harmonics. This second result is also consistent with [1] where no numerical $\log_{10} \mathcal{B}$ were reported for this comparison, since they were within numerical and systematic uncertainties – matching our results which are very close to zero.

To explore in more detail the differences between the parameters inferred using these models, we compare some of the estimated source parameters for aligned-spin vs precessing models in Fig. 3 and for dominant-mode-only vs. higher-mode models in Fig. 4. One can see here that the inclusion of precession alone does not significantly affect the posterior distributions, while much tighter constraints follow from including higher multipoles, in line with previous studies [89–91]. In Fig. 5 we also show comparisons of our own aligned-spin results against those from SEOBNRv4_ROM, NRHYBSUR3DQ8 and IMRPHENOMHM, finding excellent agreement with the first two, while IMRPHENOMHM is an outlier due to its higher modes not being directly calibrated to NR.

C. Contributions from subdominant harmonics

In order to quantify the strength of higher multipoles in GW190412, we consider the posterior samples obtained by analysing the event with IMRPHENOMXHM² (using run number

² Note that, in the twisting-up approximation, one can cleanly separate different mode-contributions in the inertial frame only in the aligned-spin limit, and

5 in Table III) and compute the SNR distribution corresponding to each mode. The optimal SNR for one mode is defined as

$$\rho_{\text{opt}}^{\text{lm}} = \sqrt{(h^{\text{lm}}|h^{\text{lm}})}, \quad (3.2)$$

where $(|)$ refers to the usual noise-weighted inner product

$$(a|b) := \text{Re} \int_0^{f_{\text{max}}} \frac{\tilde{a}(f)\tilde{b}^*(f)}{S_n(f)} df \quad (3.3)$$

and h^{lm} denotes the contribution of a given mode to the strain measured by the detector:

$$h^{\text{lm}} = F_+ h_+^{\text{lm}} + F_\times h_\times^{\text{lm}}, \quad (3.4)$$

$$h_+^{\text{lm}} - i h_\times^{\text{lm}} = h_{l m - 2} Y_{l m} + h_{l - m - 2} Y_{l - m}. \quad (3.5)$$

Fig. 6 summarizes our results. In the top panel, we compare the SNR distributions for the subdominant multipoles available in IMRPHENOMXHM, and find that the strongest contribution comes from the (3, 3) mode, in line with [1]³. In the lower panel, we compare the full optimal SNR distribution to that of the quadrupole mode only. As expected, the (2, 2) is by far the strongest mode, however, the contribution of subdominant harmonics is clearly non-negligible. This is also illustrated in Fig. 7, where we show the estimated luminosity distance and inclinations obtained when activating only the $\ell = 2$ and $\ell \leq 3$ modes and compare them with the results of a standard run, for which $\ell_{\text{max}} = 4$. One can clearly see that increasing the mode content of the model results in tighter constraints on the source location. We note also the addition of the (4, 4) mode alone, which is the second-strongest subdominant mode, has a visible effect on the posteriors.

D. Convergence, computational efficiency and cost

One of the main advantages of IMRPHENOMXPHM is its computational efficiency, which makes it a potential workhorse for large scale parameter estimation studies, and an ideal tool for convergence tests and systematic studies of sampler settings. This is particularly important considering that future sensitivity improvements of the LIGO-Virgo detector network will result in an increased detection rate [92]. The modularity of IMRPHENOMXPHM, and in particular the possibility to relax the multibanding thresholds, naturally lends itself to a hierarchical PE workflow, where computationally cheap runs can be set up to find optimal choices of priors, while more expensive runs are reserved for final data releases.

Our strategy here has been to select a default configuration with 2048 nested sampling live points (n_{live}) and ($n_{\text{act}} = 10$),

that is why we carry out this analysis with IMRPHENOMXHM.

³ We have verified that the single-mode SNR contributions calculated here are consistent with those obtained in [1], where the SNR of each subdominant mode is first orthogonalised with respect to that of the (2, 2) mode.

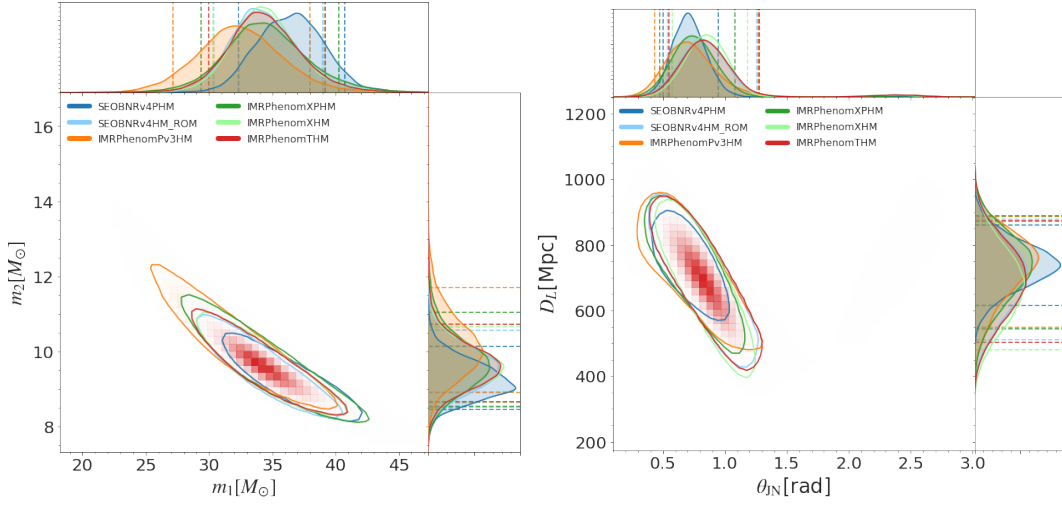


FIG. 3. Comparison of posterior distributions for the component masses, distance and inclination of GW190412, using SEOBNRv4PHM (dark blue), SEOBNRv4_ROM (light blue), IMRPHENOMPv3HM (orange), IMRPHENOMXPHM (dark green), IMRPHENOMXHM (light green) and IMRPHENOMTHM (red).

Hypotheses	Model properties	EOB	Phenom	PhenomX	PhenomT
HM vs $\ell = 2 = m $	aligned	3.5	3.4	3.5 ± 0.1	3.4 ± 0.1
	precessing	4.1	3.6	3.3 ± 0.1	-
prec. vs aligned	dominant multipoles	-	-	0.0 ± 0.1	-
	higher multipoles	-	-	0.3 ± 0.1	-

TABLE IV. Comparison of Bayes factors $\log_{10} \mathcal{B}$ between EOBNR, Phenom, PhenomX and PhenomT families for two hypothesis tests: (i) whether the signal contains higher or only dominant multipoles, (ii) whether spin precession is present or not. The multipoles test is performed for both aligned and precessing waveforms, and the test for precession both when including only dominant multipoles or also the higher multipoles. No $\log_{10} \mathcal{B}$ numbers are quoted for “prec. vs aligned” from EOBNR and Phenom because in [1] these were only noted to be not significant in comparison with statistical and systematic error bars.

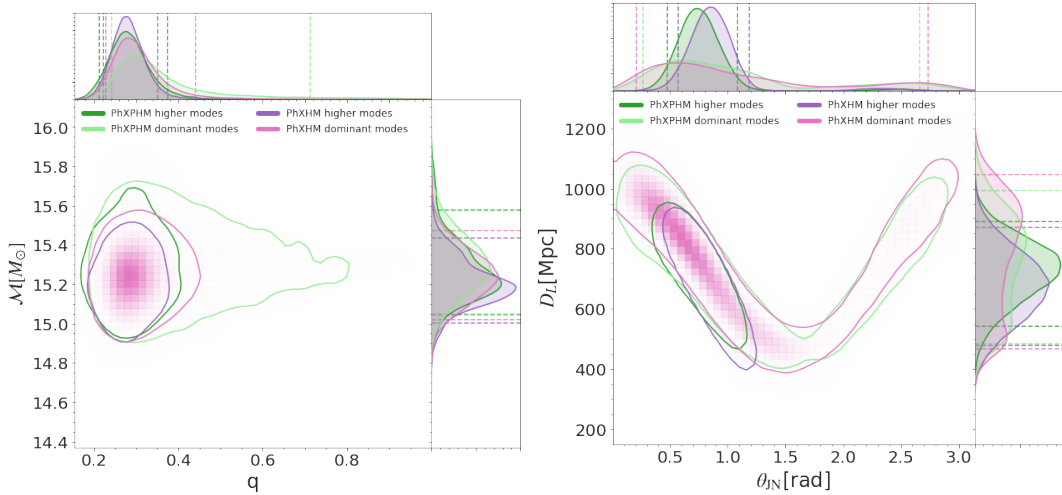


FIG. 4. Comparison of some estimated source parameters, using aligned (pink and purple) and precessing (light and dark green) waveform models from the IMRPHENOMX family. Light green and pink correspond to runs where only the quadrupole mode has been activated (note that, for the precessing model, this corresponds to the mode-content in the co-precessing frame). The inclusion of higher multipoles helps to break the degeneracy between distance and inclination, as seen in the plot on the right.

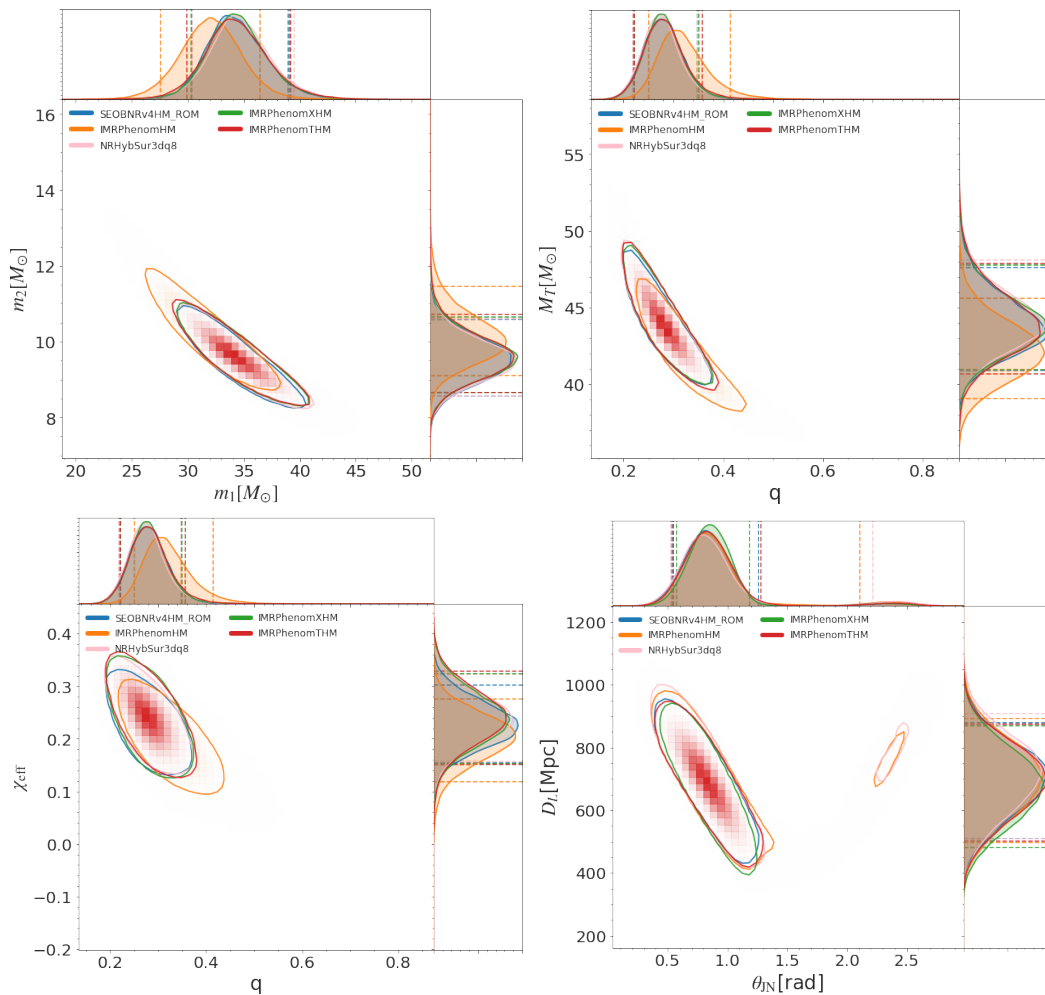


FIG. 5. Posterior distributions for the masses, spin parameters, distance and inclination of GW190412, as estimated with aligned-spin waveform models (SEOBnrV4_ROM: blue, NRHybSur3dq8: pink, IMRPHENOMHM: orange, IMRPHENOMXHM: green, IMRPHENOMTHM: red). Models with multipoles individually calibrated to NR (SEOBnrV4_ROM, IMRPHENOMXHM and IMRPHENOMTHM) are in excellent agreement with the numerical relativity surrogate NRHybSur3dq8, while IMRPHENOMHM returns visibly shifted posteriors.

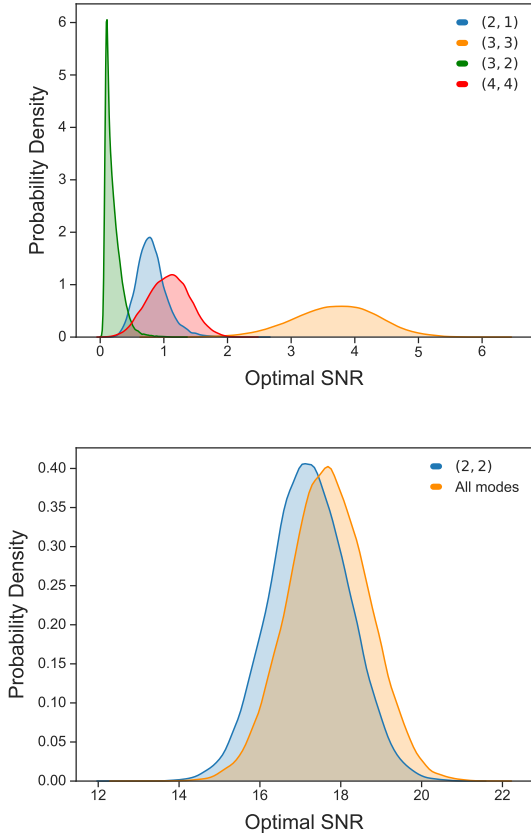


FIG. 6. Top panel: Optimal SNR distributions for higher multipoles, when recovering the signal with IMRP_{PHENOMX}HM. The waveform used to compute the SNR contains just one single mode. Bottom panel: SNR distribution for the dominant (2, 2) mode only waveform and the whole multimode waveform.

corresponding to run 23 in Table III.⁴ As we report below, from all our tests this configuration gives robust results when compared both with more expensive settings (increasing n_{live} and n_{act}), and with computationally cheaper settings. This justifies to use our default setting for comparisons of different choices of precession treatment in appendices C and D, and for comparisons with runs that reduce the physics content in the model to non-precessing spins or the dominant quadrupole in Sec. III B. We will also show here that configurations with drastically reduced computational cost still reproduce the posteriors of more accurate runs rather well, and are sufficient for fast exploratory runs, e.g. to determine prior settings, or to obtain rapid accurate distance measurements that take advantage of precession and higher harmonics.

Table III summarizes all the runs performed with parallel Bilby and distance marginalisation switched on, with differ-

⁴ Run 26 with $n_{\text{act}} = 50$ has been used for our main astrophysical results instead, but run 23 with $n_{\text{act}} = 10$ is the baseline for efficiently comparing to alternative settings and models.

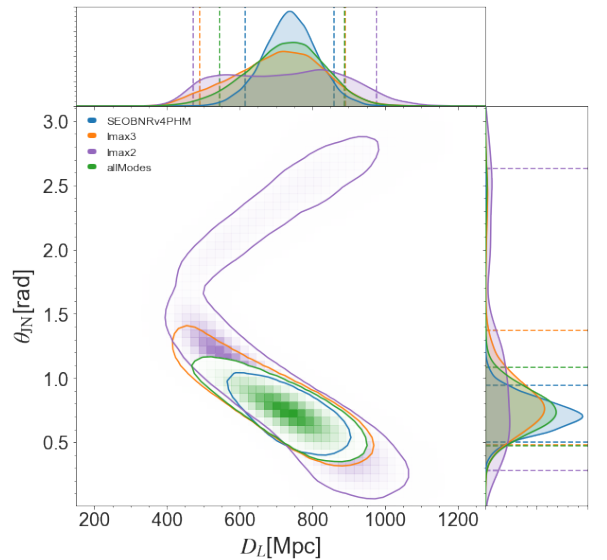


FIG. 7. Joint posterior distributions for luminosity distance and inclination of GW190412, as estimated with SEOBNRv4PHM (blue), IMRP_{PHENOMX}PHM with its full mode-content (green) and IMRP_{PHENOMX}PHM with only the $\ell = 2$ (purple) or $\ell \leq 3$ modes activated.

ent models of the IMRP_{PHENOMX} family as well as IMRP_{PHENOMX}THM. Additional parallel Bilby runs without distance marginalisation are listed in Table VIII. Most of our runs were performed on IntelXeon Platinum CPUs with 2.1GHz clock-rate (BSC MareNostrum), except for some runs in Table VIII, which have been performed on Intel Xeon E5-2670 CPUs with 2.6GHz clock-rate (Picasso machine). All the runs of Table III used the master branch of LALSuite, with git hash f253e1307b9c19b0fa974fe627651db483f38170, compiled with gcc version 5.4.0 (BSC MareNostrum) and 4.9.4 (Picasso). The cost of a full higher-mode preprocessing run with $n_{\text{live}} = 2048, n_{\text{act}} = 10$ can be as low as ≈ 1300 cpu hrs per seed, using aggressive multibanding thresholds and distance marginalization, which translates into a total sampling time of roughly 13 hrs when using 96 CPU-cores. The cost can be further decreased by lowering the number of live points: the fastest run in our analysis, using $n_{\text{live}} = 512, n_{\text{act}} = 10$ had a computational cost of less than 700 CPU h per seed. We should stress the fact that the computational costs reported here correspond to the analysis of 8 s of data and will be therefore even lower for BBH events with high total masses, for which segment lengths of 4 s are appropriate.

Fig. 8 compares some estimated source parameters obtained with IMRP_{PHENOMX}PHM, for different combinations of multibanding thresholds and sampler settings. Even for the most aggressive settings, with moderately high multibanding thresholds and $n_{\text{live}} = 512$, one cannot appreciate large differences with respect to our default run.

To quantify the robustness of results under changes of sampler settings, we have computed the JS divergences (see definition in Sec. II D) between IMRP_{PHENOMX}PHM runs employing

different numbers of n_{live} and/or n_{act} . Our results are summarised in Table VI. For comparison Table V lists JS divergences between runs with different models.

Among the divergences of runs with different sampler settings but the same IMRP_{HENOM}XPHM model, statistically significant discrepancies (defined as $JS \geq 0.002$, according to the same criterion adopted in [47]), occur more frequently for the cheapest configurations ($n_{\text{live}} = 512$), as expected. We also observe that $n_{\text{live}} = 2048, n_{\text{act}} = 10$ gives already entirely satisfactory results when compared with more expensive settings and could be therefore considered a safe configuration for production runs on this event.

XAS	-			
XHM	$JS_{\phi_c} = 0.2929$	-		
XP	$JS_{t_1} = 0.6689$	$JS_{t_1} = 0.6863$	-	
XPHM	$JS_{t_1} = 0.6691$	$JS_{t_1} = 0.68791$	$JS_{\psi} = 0.2659$	-
	XAS	XHM	XP	XPHM

TABLE V. Maximum Jensen-Shannon (JS) divergence values between the posterior distributions of GW190412 estimated with IMRP_{HENOM}XAS, IMRP_{HENOM}XHM, IMRP_{HENOM}XP and IMRP_{HENOM}XPHM.

IV. FURTHER WAVEFORM SYSTEMATICS STUDIES

A. Matches for low total-mass binaries

Waveform systematics can be further investigated by computing noise weighted frequency-domain overlaps between model and signal waveforms:

$$\langle h_M, h_S \rangle = 4\Re \int_{f_{\min}}^{f_{\max}} \frac{\tilde{h}_M(f) \tilde{h}_S^*(f)}{S_n(f)}, \quad (4.1)$$

where $S_n(f)$ is the one-sided PSD of the detector noise. In what follows, we will take IMRP_{HENOM}XPHM as the model and assume that the signal is given either by a NR or SEOBNRv4PHM waveform. We will quantify the agreement between model and signal by means of the match function:

$$\mathcal{M}(h_M, h_S) = \max_{t_c, \phi_0, \psi_0} \frac{\langle h_M, h_S \rangle}{\sqrt{\langle h_M, h_M \rangle} \sqrt{\langle h_S, h_S \rangle}}, \quad (4.2)$$

where we optimize the overlap between normalized waveforms over polarization angle ψ_0 , coalescence time t_c and reference phase ϕ_0 . In general, the definition of the spin configuration at the reference frequency will not be the same for the model and the signal. To account for this, we also optimize the match over rotations of the initial in-plane spins. We use the Advanced-LIGO [5] design sensitivity Zero-Detuned-High-Power PSD [93] with $f_{\min} = 20$ Hz and $f_{\max} = 2048$ Hz.

Match calculations are far cheaper than a fully fledged Bayesian analysis and can be leveraged for extensive explorations in parameter space. We extended our previous results

[10] and specifically targeted low total-mass binaries similar to GW190412.

We computed the match between SEOBNRv4PHM and IMRP_{HENOM}XPHM waveforms for 20000 random configurations with mass ratios uniformly distributed in the interval $q \in [1, 5]$, spins isotropically distributed on the unit sphere and total mass varying between $30M_{\odot}$ and $50M_{\odot}$ in bins of $5M_{\odot}$ width. Our results are illustrated in Fig. 9. The top panel shows the normalized probability distribution of $\log(1 - \mathcal{M})$. Overall, the two models agree very well, with $\approx 90\%$ of the configurations achieving $\mathcal{M} \geq 95\%$; however, it can be seen that matches tend to degrade for higher mass ratios. The lower panel shows instead the match \mathcal{M} as a function of the mass ratio q and the effective precession spin χ_p . We notice that the two models can show fairly large disagreement when it comes to strongly precessing, high mass ratio systems.

We have also checked the agreement between the model and NR for low total mass binaries. The NR simulations considered here are all public SXS precessing waveforms available in the `1vcnr` catalog [94]. For simplicity, we focused on systems with $q \in [2, 5]$, as all the precessing higher mode models considered here tend to exclude mass ratios outside of this range for GW190412. In this case we quantify the agreement between model and signal (NR) through the SNR-weighted match \mathcal{M}_w [95]

$$\mathcal{M}_w = \left(\frac{\sum_i \mathcal{M}_i^3 \langle h_{i,\text{NR}}, h_{i,\text{NR}} \rangle^{3/2}}{\sum_i \langle h_{i,\text{NR}}, h_{i,\text{NR}} \rangle^{3/2}} \right)^{1/3}, \quad (4.3)$$

where the subscript i goes over different polarization and reference phases of the signal. Fig. 10 shows our results: coloured lines correspond to simulations for which the maximum mismatch $1 - \mathcal{M}_w$ is above 3% in the mass range considered. The three outliers with mismatches above this threshold are SXS:0057, SXS:0059 and SXS:0062, which all correspond to $q = 5$ binaries. These result confirm the conclusions drawn in [10]. Overall, SEOBNRv4PHM and IMRP_{HENOM}XPHM show good agreement over most of the parameter space and are expected to deliver comparable estimates for typical BBH events. However, exceptional events with very asymmetric masses and/or strong precession might require a more in-depth analysis, which we leave for future work. Further insight on waveform systematics can be found in previous works on the IMRP_{HENOM}X family models [7–10] where extensive comparisons through matches with previous waveform models and NR simulations are carried out.

B. Injection study

In order to assess the accuracy of our measurements, in particular of the precessing spin, we injected into a Hanford-Livingston-Virgo detector network three synthetic signals and recovered their parameters with IMRP_{HENOM}XPHM. Two of these signals are based on NR simulations with moderately high χ_p , while the third one makes use of a SEOBNRv4PHM

N512 NA10	-							
N512 NA50	$JS_{t_1} = 0.0010$	-						
N1024 NA10	$JS_{t_2} = 0.0032$	$JS_{t_2} = 0.0023$	-					
N1024 NA50	$JS_{t_1} = 0.0008$	$JS_{ra} = 0.0007$	$JS_{ra} = 0.0019$	-				
N2048 NA10	$JS_{ra} = 0.0020$	$JS_{ra} = 0.0020$	$JS_{t_2} = 0.0010$	$JS_{ra} = 0.0017$	-			
N2048 NA50	$JS_{\phi_{JL}} = 0.0015$	$JS_{\phi_{JL}} = 0.0016$	$JS_{t_2} = 0.0014$	$JS_{a_2} = 0.0017$	$JS_{\phi_{JL}} = 0.0013$	-		
N4096 NA10	$JS_{ra} = 0.0023$	$JS_{ra} = 0.0027$	$JS_{\theta_{JN}} = 0.0030$	$JS_{\theta_{JN}} = 0.0018$	$JS_{\theta_{JN}} = 0.0033$	$JS_{\theta_{JN}} = 0.0020$	-	
	N512 NA10	N512 NA50	N1024 NA10	N1024 NA50	N2048 NA10	N2048 NA50	N4096 NA10	

TABLE VI. Maximum Jensen-Shannon (JS) divergence values between the posterior distributions of GW190412 estimated with IMRP_{HENOM}XPHM using different sampler settings. In this table, N is short for n_{iive} and NA is short for n_{act} . We report here only values larger than 0.001.

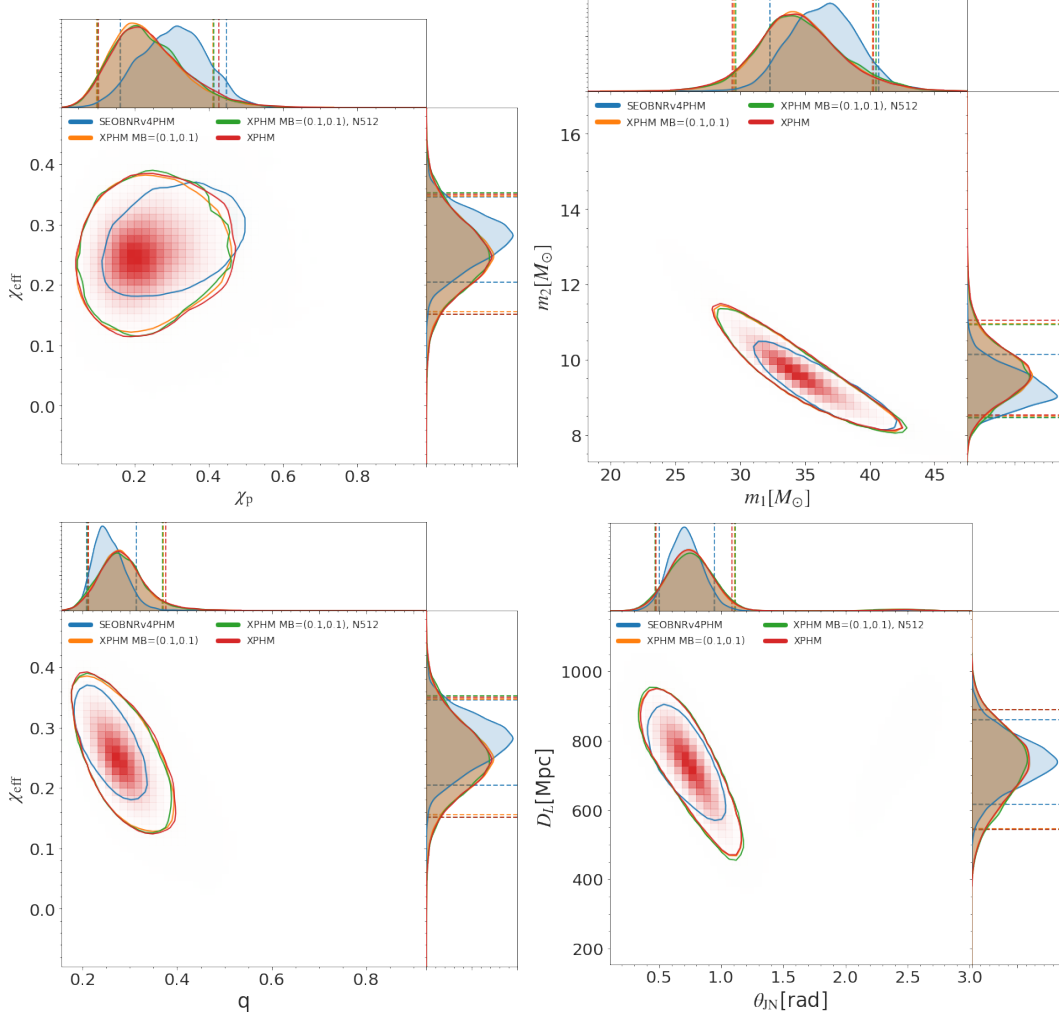


FIG. 8. Comparison of some estimated source parameters obtained with SEOBNRv4PHM (blue) and IMRP_{HENOM}XPHM, with different combinations of multibanding thresholds and sampler settings (red for the default version, orange with aggressive multibanding thresholds for both the aligned-spin waveforms and precession angles, green with aggressive multibanding and $n_{\text{iive}} = 512$).

waveform, corresponding to the maximum-likelihood sample of the LVC release. We injected the signals in zero noise, but we fed into the likelihood estimation the PSDs used to analyse GW190412. For the NR signals, the total mass and luminosity distance were chosen so as to achieve a network

SNR comparable to that of GW190412. In all cases, we set the reference frequency to be the same as the minimum frequency, i.e. $f_{\text{ref}} = f_{\text{low}}$ in our setup.

The first injection made use of the public NR waveform SXS:BBH:0049, corresponding to a $q = 3$ binary with $\chi_p \approx 0.5$.

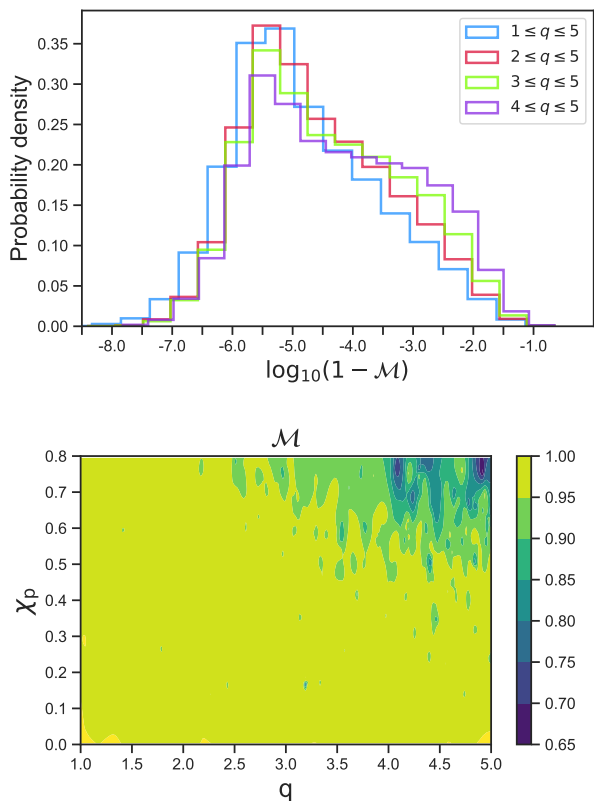


FIG. 9. Matches between IMRPHENOMXPHM and SEOBNRv4PHM waveforms. The top panel shows the distribution of mismatches for different ranges in mass-ratio q . The lower panel shows the match as a function of the mass ratio q and effective precession spin χ_p . Light green (dark blue) areas indicate good (poor) agreement between the two models. The worst matches correspond to strongly precessing, high mass-ratio binaries.

We simulated a binary with right ascension $ra = 3.81$ rad, declination $dec = 0.63$ rad, and geocentric time equal to the trigger time reported in GraceDB [80]. We used a minimum frequency of 25.7 Hz, due to the limited length of the NR waveform. In Fig. 11 we present posteriors for detector-frame masses, χ_{eff} , χ_p , luminosity distance, θ_{JN} , as well as the individual adimensional spin magnitudes a_1 and a_2 . Overall, we observe a very good agreement between the recovered parameters and the injected values, which always lie within the 90% credible intervals of the posterior distributions. In particular, the spin magnitude of the primary is very well constrained, similarly to what happens for GW190412.

In Fig. 12 we present the results of our second NR injection, that took as input the public SXS waveform SXS:BBH0058, corresponding to a $q = 5$ binary with $\chi_p \approx 0.5$. The simulated event was assumed to have the same celestial coordinates and geocentric time as the previous mock signal. The total mass of the system was taken to be $50.5 M_\odot$ in the detector frame, and we used a minimum frequency of 20.5 Hz. We can see that, even in this case, all the recovered parameters are consistent

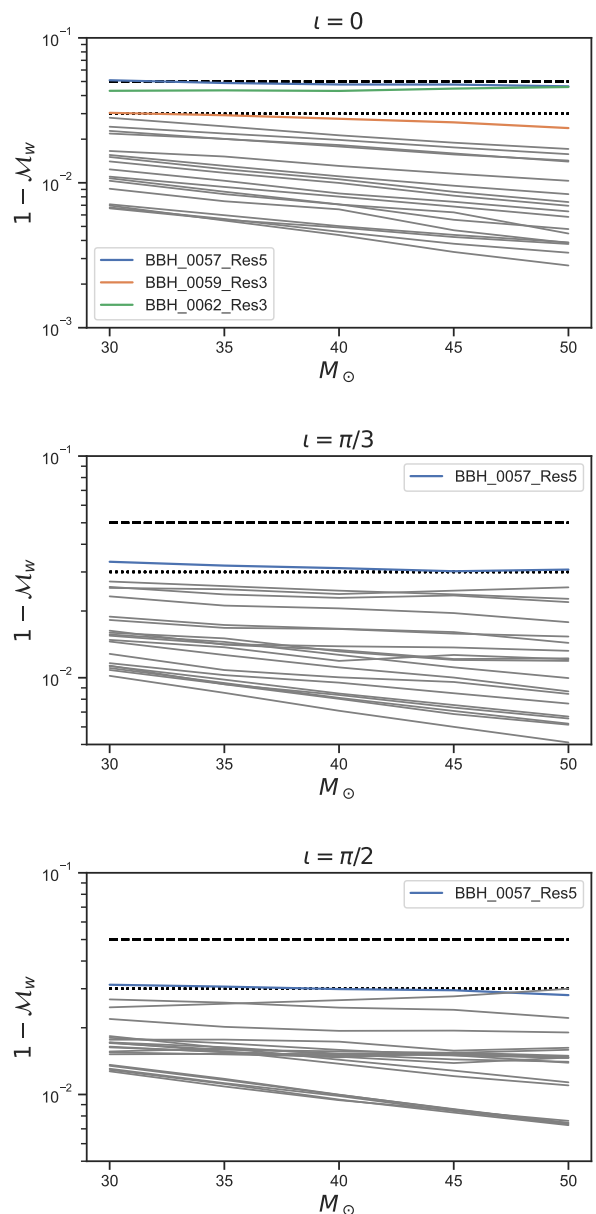


FIG. 10. SNR-weighted mismatches between IMRPHENOMXPHM and precessing SXS waveforms with $q \in [2, 5]$. The dotted and dashed lines in each panel denote thresholds of 3% and 5%, respectively. Cases for which the maximum mismatch exceeds the 3% threshold are shown with coloured lines and listed in the legends.

with the ones of the injected signal.

Finally, we injected in zero-noise the maximum-likelihood waveform for the SEOBNRv4PHM sample. The injected signal had a total mass of $46.66 M_\odot$ in detector frame, celestial coordinates $ra \approx 3.81$ rad and $dec \approx 0.64$ rad, and geocentric time of approximately 1239082262.18 s. Fig. 13 summarizes our main results. IMRPHENOMXPHM recovers well all the parameters of the injected signal. We therefore conclude that

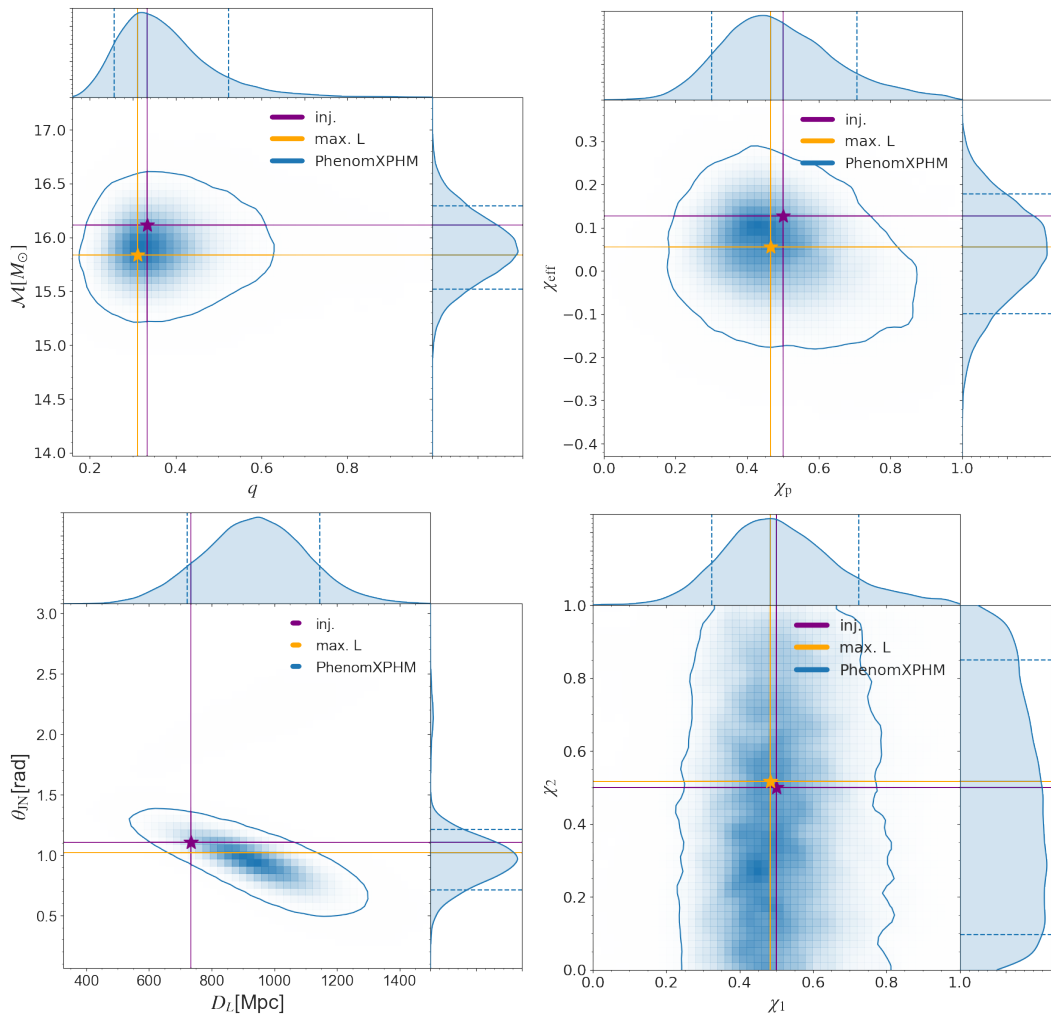


FIG. 11. Posterior distributions for some of the intrinsic and extrinsic parameters of a synthetic signal obtained starting from the public SXS waveform SXS:BBH0049. Purple and orange stars mark the values of the injected and maximum-likelihood parameters, respectively. Dashed lines indicate the 90% credible intervals of the individual posterior distributions.

IMR_{PHENOMXPHM} appears to deliver robust estimates of the source properties for events similar to GW190412.

V. CONCLUSIONS

We have re-analysed the event GW190412 with the newest generation of phenomenological waveform models: the frequency-domain model IMR_{PHENOMXPHM}, which includes precession and higher modes, and the time-domain model IMR_{PHENOMTHM}, which includes higher modes, but not yet precession. Both models have been constructed with similar techniques and accuracy goals, and we refer to them (and several versions with reduced physics content) jointly as “generation X”.

The principal code we have used for performing Bayesian inference is parallel Bilby [46], which allows us to work in traditional high performance computing environments and to

obtain results on a time scale of hours. In Sec. III D we have described some tests varying sampling parameters to convince us that our parallel Bilby runs have converged to reliable results. In appendix B we further test that our main runs, which use distance marginalisation, agree with runs that do not use it, and we study the difference in the number of likelihood evaluations and computational cost. Since parallel Bilby is a relatively new code, we also report our cross-checks with the more established LALInference code [48] in appendix A, and find excellent agreement.

For the non-precessing sector, we find excellent agreement between all waveform models where sub-dominant harmonics have been calibrated to numerical waveforms, i.e. SEOBNRv4HM, NRHybSur3dq8, IMR_{PHENOMXHM} and IMR_{PHENOMTHM} (see e.g. Fig. 5). However we find significant differences for IMR_{PHENOMHM}, where higher modes have been constructed by an approximate map in terms of the dominant quadrupole mode, which was calibrated to a much smaller set of

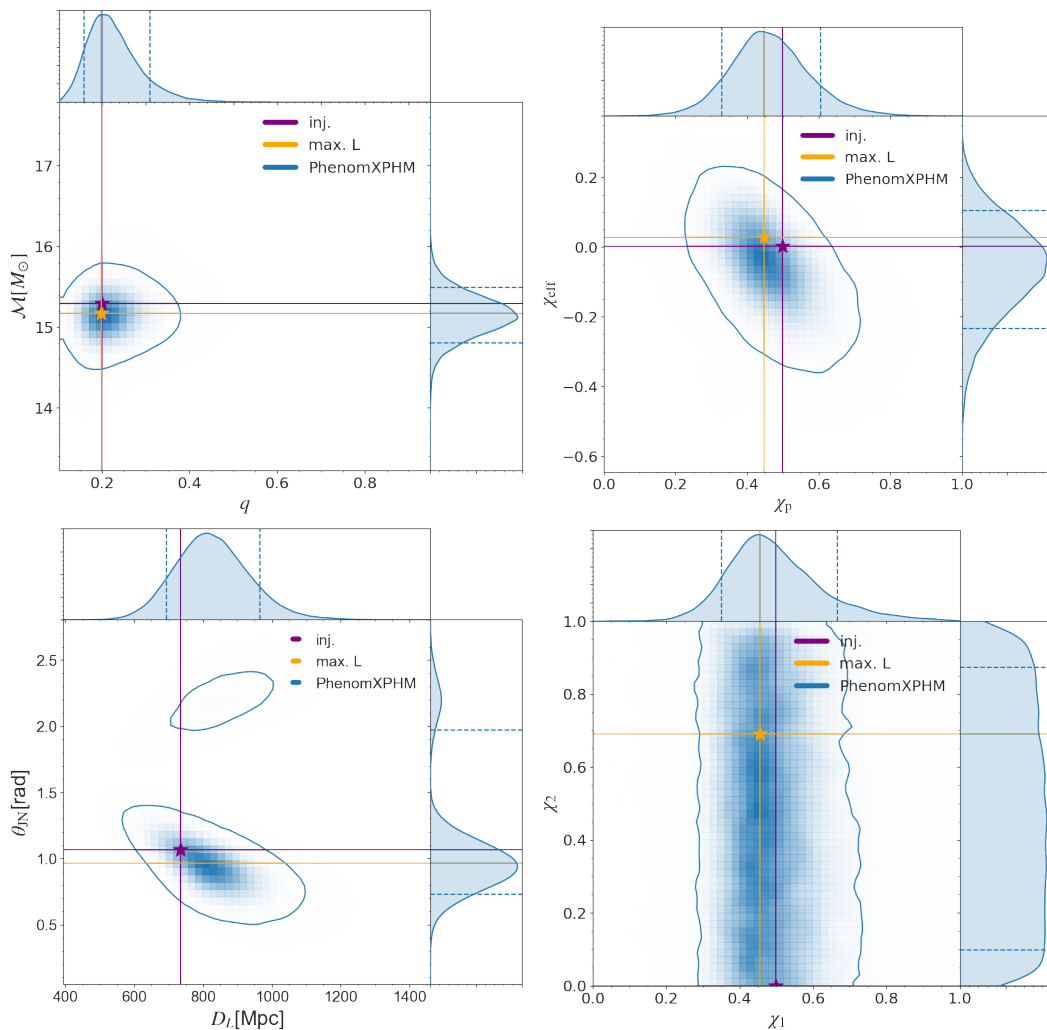


FIG. 12. Posterior distributions for some of the intrinsic and extrinsic parameters of a synthetic signal obtained starting from the public SXS waveform SXS:BBH0058. Purple and orange stars mark the values of the injected and maximum-likelihood parameters, respectively. Dashed lines indicate the 90% credible intervals of the individual posterior distributions.

NR waveforms than IMR_{PHENOMXAS}. It is to be expected that in the future, EOB and phenomenological models will also be calibrated to numerical relativity waveforms for the precessing sector, and will thus reach similar levels of agreement.

When adding precession we find that the agreement between phenomenological waveforms and the EOB model improves for some quantities, notably the masses and source location. Not surprisingly, for the effective precession spin this is not the case, although the posterior width for IMR_{PHENOMXPHM} is closer to SEOBNRv4PHM than IMR_{PHENOMPv3HM}. Both SEOBNRv4PHM and the frequency-domain phenomenological waveforms use a version of the “twisting-up” approximate map between non-precessing and precessing waveforms [37], however the current frequency domain phenomenological waveforms employ the stationary phase approximation (SPA) in addition to the approximations inherent in the twisting-up (for a recent discussion of these approximations see [96]). An improved accuracy for precession is thus expected from the

extension of the LALSuite implementation of IMR_{PHENOMT} to include precession following [11]. Strategies for extending the twisting-up procedure for frequency-domain waveforms beyond the SPA approximation have been discussed in [97]. In appendix C we compare the two different descriptions of the Euler angles used in the twisting-up procedure implemented in IMR_{PHENOMXPHM} (single-spin and double-spin), and only find small differences, mainly in the effective precession spin parameter χ_p . Not surprisingly we find that the double-spin description is closer to the results that have been reported for SEOBNRv4PHM in [15]. In the same section we also compare different prescriptions for the spin of the merger remnant in precessing mergers, and we find no significant differences for GW190412. In appendix D we compare directly against an alternative “twisting-up” of the older IMR_{PHENOMHM} aligned-spin model implemented within the IMR_{PHENOMXPHM} framework, finding consistent results with the original IMR_{PHENOMPv3HM} (which is based on IMR_{PHENOMHM}), thus cleanly demonstrat-

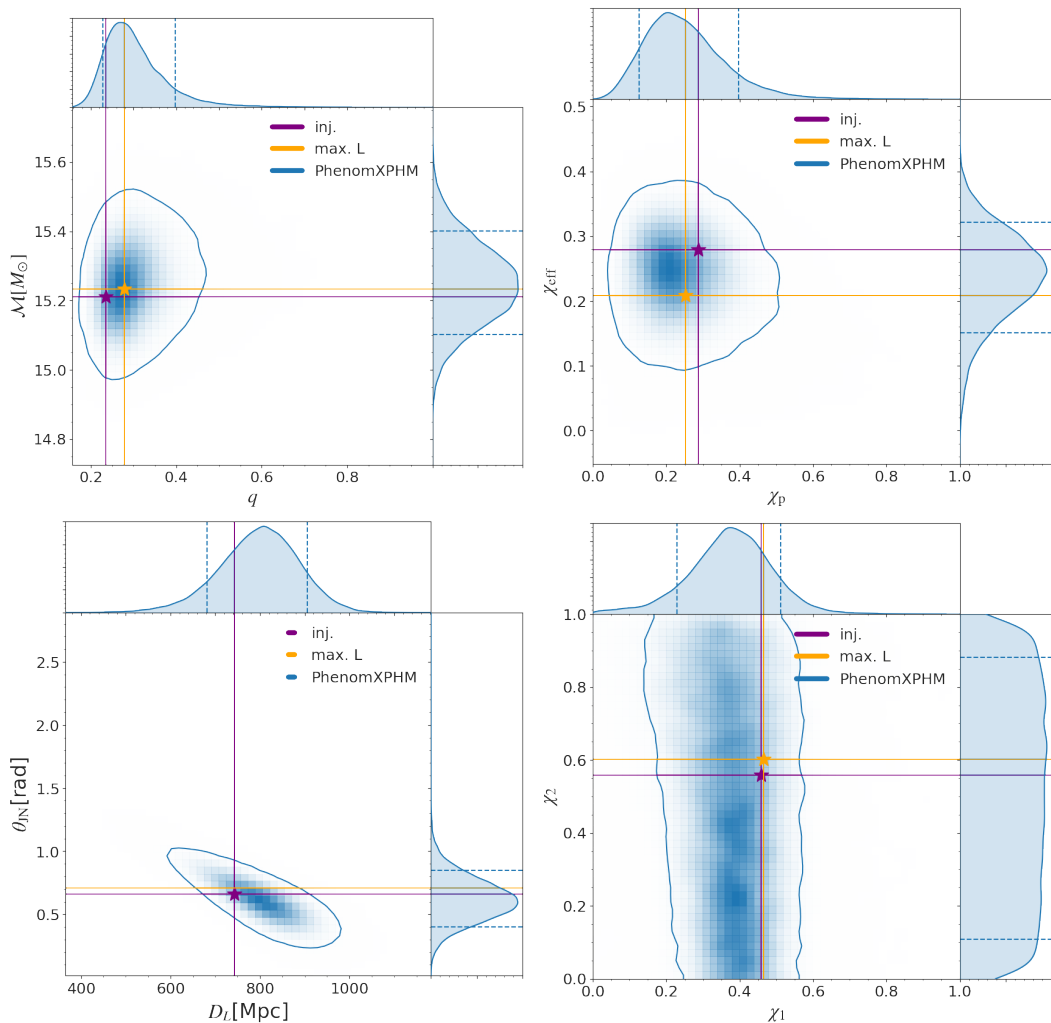


FIG. 13. Posterior distributions for some of the intrinsic and extrinsic parameters of a synthetic signal obtained from the SEOBNRv4PHM maximum-likelihood parameters for GW190412. Purple and orange stars mark the values of the injected and maximum-likelihood parameters, respectively. Dashed lines indicate the 90% credible intervals of the individual posterior distributions.

ing that our improved results mostly derive from the updated underlying aligned-spin model IMRPHENOMXHM with its calibration of subdominant modes to NR. We have also shown that the use of the multibanding algorithm leads to a dramatic reduction of the computational cost of parameter estimation runs, nearly halving the cost of non-multiband runs.

Additional studies presented in this paper include investigations into possible remaining systematic differences between waveforms, including waveform match comparisons and injection studies in Sec. IV. The tested injections include SXS waveforms with parameters consistent with GW190412 as well as the SEOBNRv4PHM maximum-likelihood waveform, finding no significant systematic issues with IMRPHENOMXPHM in this part of the parameter space. We also test the impact of different spin priors in appendix E, finding that PE results for GW190412 are generally consistent when changing from the standard LVC prior to a volumetric one, and confirming the results of [58] that the standard assumption of allowing for a

spinning heavier BH component is preferred over runs with a prior that restricts spin to the less massive BH only.

One of the key properties of the IMRPHENOMX waveform family is its computational efficiency. We demonstrate that even the most complete IMRPHENOMXPHM model, when making use of the pBilby sampler and a strong HPC cluster, allows for extremely fast-turn-around exploratory runs (few hours) and still very fast high-fidelity runs like the ones we report as the main results (see e.g. Fig. 1, with the plotted IMRPHENOMXPHM run taking 2670 CPU hours total and finishing after 28 hours wall-clock time). The time-domain IMRPHENOMTHM model is also quite competitive in run time despite being a native time domain model, having only 50 more CPU cost per likelihood evaluation than IMRPHENOMXHM. More details on computational cost for different model versions and sampler settings are found in Table III (for the runs with distance marginalization) and Table VIII (for the runs without distance marginalization). As expected, the use of distance marginal-

ization brings significant computational cost savings: this can be readily appreciated by comparing otherwise equivalent runs from the two tables (e.g. run 23 in Table III and run 20 in Table VIII).

In summary, the studies presented here demonstrate the robustness and efficiency of the “generation X” phenomenological waveform models in a real-world application, and point out how systematic runs with varied settings can be used both to study the physics of an individual detection in detail, and to achieve accuracy requirements at bounded computational cost. We hope these results will help to advance the routine use of subdominant harmonics in the parameter estimation for compact binary mergers, leading to both more accurate parameter estimates, and the availability of accurate posterior estimates on the time scale of a few hours.

Posterior samples from our preferred run for each waveform model are released in [49].

ACKNOWLEDGEMENTS

We gratefully thank the Bilby and parallel Bilby code developers, especially Greg Ashton, Sylvia Biscoveanu, Rory Smith and Colm Talbot for discussions and software fixes. This work was supported by European Union FEDER funds, the Ministry of Science, Innovation and Universities and the Spanish Agencia Estatal de Investigación grants PID2019-106416GB-I00/AEI/10.13039/501100011033, FPA2016-76821-P, RED2018-102661-T, RED2018-102573-E, FPA2017-90687-REDC, Vicepresidència i Conselleria d’Innovació, Recerca i Turisme, Conselleria d’Educació, i Universitats del Govern de les Illes Balears i Fons Social Europeu, Comunitat Autònoma de les Illes Balears through the Direcció General de Política Universitària i Recerca with funds from the Tourist Stay Tax Law ITS 2017-006 (PRD2018/24), Generalitat Valenciana (PROMETEO/2019/071), EU COST Actions CA18108, CA17137, CA16214, and CA16104, and the Spanish Ministry of Education, Culture and Sport grants FPU15/03344 and FPU15/01319. M.C. acknowledges funding from the European Union’s Horizon 2020 research and innovation programme, under the Marie Skłodowska-Curie grant agreement No. 751492. D.K. is supported by the Spanish Ministerio de Ciencia, Innovación y Universidades (ref. BEAGAL 18/00148) and cofinanced by the Universitat de les Illes Balears. The authors thankfully acknowledge the computer resources at MareNostrum and the technical support provided by Barcelona Supercomputing Center (BSC) through Grants No. AECT-2019-2-0010, AECT-2019-1-0022, from the Red Española de Supercomputación (RES). Authors also acknowledge the computational resources at the cluster CIT provided by LIGO Laboratory and supported by National Science Foundation Grants PHY-0757058 and PHY-0823459. This research has made use of data obtained from the Gravitational Wave Open Science Center [98], a service of LIGO Laboratory, the LIGO Scientific Collaboration and the Virgo Collaboration. LIGO is funded by the U.S. National Science Foundation. Virgo is funded by the

French Centre National de Recherche Scientifique (CNRS), the Italian Istituto Nazionale della Fisica Nucleare (INFN) and the Dutch Nikhef, with contributions by Polish and Hungarian institutes. This work has been assigned LIGO document number P2000402.

Appendix A: Comparison between pBilby and LALInference

In order to cross-validate our parallel Bilby results, we perform the runs listed in Table VII using the Bayesian parameter estimation package LALInference (LI) [48]. The runs that make use of nested sampling employ five different seeds, 2048 live points, and a maximum chain length of 5000. Those that use MCMC sampling, utilize eight parallel tempering temperatures and 24 parallel chains.

In Fig. 14 we compare the standard runs of Table III using $n_{\text{live}} = 2048$ and $n_{\text{act}} = 10$ (dashed) with the default LI runs (solid). All posteriors are well recovered, and we find generally good agreement between pBilby and LI runs. We also observe that the agreement is better in the presence of higher modes. A possible explanation is that the breaking of degeneracies due to higher modes reduces the posterior volume and thus benefits better sampling.

Appendix B: Comparison of runs with and without distance marginalisation

We report in Table VIII the complete list of pBilby runs performed on GW190412 without using distance marginalisation (prior to a bugfix in the distance marginalisation algorithm). We find excellent agreement and show a comparison in Fig. 15 for some key quantities.

Appendix C: Comparison of IMRP_{PHENOMX} precession and final spin versions

We have performed several IMRP_{PHENOMXPHM} runs with non-default waveform options to study the robustness of our results under changes of the precession prescription and final spin version. We compare posterior distributions for the default and non-default runs in Fig. 16. Mass parameters appear to be insensitive to changing these settings, while we observe minor differences in the spin parameters, with the MSA prescription returning slightly broader posteriors.

Appendix D: Comparison with IMRP_{PHENOMPV3HM}

As we observed in the previous subsection, there are visible differences in the posterior distributions obtained with IMRP_{PHENOMPV3HM} and IMRP_{PHENOMXPHM}. We mean to quantify here the effect of the underlying aligned-spin model

Approximant	Run n	Modes (l,m)	PV FS PMB MB	Prior	Sampler	nlive	max. mcmc	nparallel	ntemps
IMRPhenomXAS	1	(2,2)	- - - -	Aligned spin	Nested	2048	5000	5	-
IMRPhenomXHM	2	D	- - - D	Aligned spin	Nested	2048	5000	5	-
IMRPhenomXP	3	(2,2)	D D D D	Precessing	Nested	2048	5000	5	-
	4	(2,2)	102 D D D	Precessing	Nested	2048	5000	5	-
IMRPhenomXPHM	5	D	D D D D	Precessing	MCMC	-	-	24	8
	6	D	102 D D D	Precessing	MCMC	-	-	24	8
	7	D	102 D D D	Precessing	Nested	2048	5000	5	-

TABLE VII. List of all LALInference runs we performed on open data for GW190412 using the IMRPHENOMX family. For each run, we indicate the LALSuite waveform approximant called along with along with various waveform settings: the mode content, precessing prescription (PV), final spin version (FS), and multibanding thresholds applied to the evaluation of the Euler angles (PMB) and of the aligned-spin modes (MB); as well as the prior used. We also denote the sampler used (Nested or MCMC) and its corresponding settings such as number of live points (nlive), maximum chain length (max. mcmc), parallel chains (nparallel) and the number of parallel tempering temperatures (ntemps).

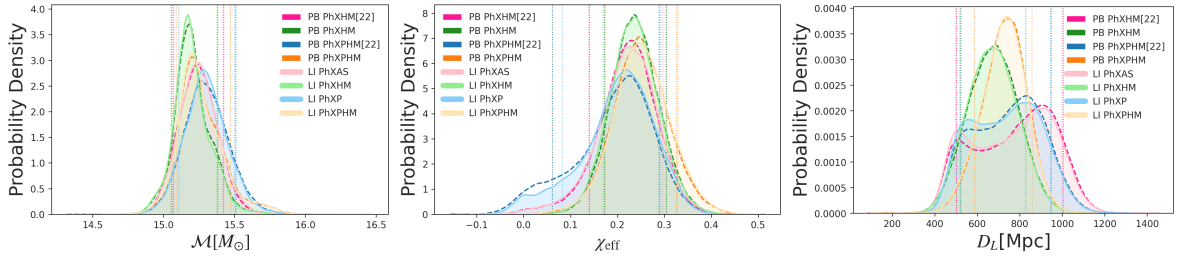


FIG. 14. Posterior distributions of several source parameters comparing the LALInference (LI) runs as listed in Table VII (solid) and the standard parallel Bilby runs from Table III using $n_{\text{live}} = 2048$ and $n_{\text{act}} = 10$ (dashed).

on those results. This can be easily done with IMRPHENOMX-PHM, which has an in-built option allowing to twist-up the GW modes returned by the older model IMRPHENOMHM (we will refer to this configuration as “IMRPhenomPHM”) instead of IMRPHENOMXHM. Using this option, one can cleanly separate the effects of the HM calibration from all the details of the precessing extension. Fig. 17 compares previous results with the IMRPhenomPHM run described above.

It can be seen that the shifts observed in the posteriors can be ascribed to the different underlying aligned-spin models, as, despite the use of independent precessing extensions, IMRPHENOMPV3HM and IMRPhenomPHM return equivalent results. Comparing both against the default IMRPHENOMXPHM, we observe in particular a marked difference in the recovery of the geocentric time of the event, due to the improved time-alignment provided by the IMRPHENOMX models.

Appendix E: Investigating different spin priors

For the vast majority of our runs we adopted the same priors employed in the LVC analysis [1]. Specifically, we choose uniform priors for detector-frame mass ratio $0.125 \leq q \leq 1$ and chirp mass $13 \leq \mathcal{M} \leq 18$, with masses constrained to lie in the interval $5 \leq m_{1,2} \leq 60$. We use a power law prior with exponent 2 for the luminosity distance, with a lower bound of 100 Mpc and uniform priors for phase and polarization angle. We also investigate the effect of different spin priors on the posterior es-

timates, in particular on the χ_p parameter whose interpretation is subtle. Our default prior for precessing runs⁵, in keeping with the LVC standard (defined as prior “P1” in appendix C.1 of [14] and also used in the analysis of GW190412 [1]), employs uniform distributions in the dimensionless component spin magnitudes ($a_i \in [0.0, 0.99]$) and isotropically distributed spin tilts (uniform in $\cos(\theta_i)$). Another common prior choice is a “volumetric spin” prior, where the magnitudes follow a power law with exponent 2 (prior “P2” in appendix C.1 of [14]).

As shown in Fig. 18, the induced prior on χ_p from the volumetric prior prefers higher values than for the standard prior. We compare two IMRPHENOMXPHM runs with these two prior choices, both with 2048 live points, $n_{\text{act}} = 5$ and without distance marginalisation (runs 12 and 19 in Table VIII). We find a slight shift of the inferred χ_p towards higher values: $\chi_p = 0.26^{+0.19}_{-0.12}$ from the run with a volumetric prior vs. $\chi_p = 0.22^{+0.21}_{-0.13}$ from the run with the default prior. However the resulting estimate is still lower than those reported in [1] (e.g. $\chi_p = 0.31^{+0.14}_{-0.15}$ for SEOBNRv4PHM). Other relevant parameters like χ_{eff} show no noticeable change under this prior.

While the Bayes factor for precession against the corresponding IMRPHENOMXHM run is indecisive for the default prior ($\Delta \log_{10} B = 0.06$), for the volumetric prior precession is actually slightly disfavoured ($\Delta \log_{10} B = -0.84$). Since the volu-

⁵ For aligned-spin runs we use instead a ‘z-prior’, corresponding to a projection of the isotropic spin prior along the direction of the total angular momentum.

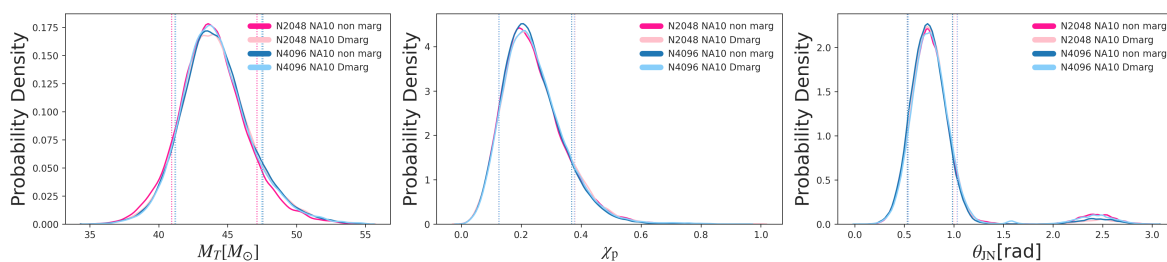


FIG. 15. Comparison between some key posterior distributions obtained with and without distance marginalisation, using two sets of sampler settings ($n_{\text{live}} = 2048$ and $n_{\text{act}} = 10$ vs. $n_{\text{live}} = 4096$ and $n_{\text{act}} = 10$).

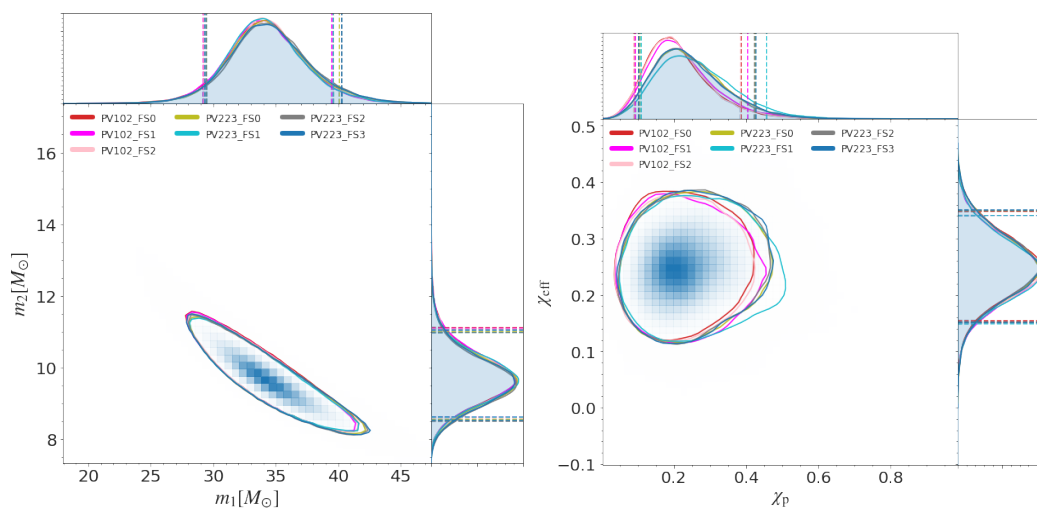


FIG. 16. 1D and 2D posterior distributions for detector-frame masses and spin parameters of GW190412 obtained with three versions of IMRPHENOMXPHM that implement different precessing prescriptions (102 for NNLO angles and 223 for MSA) and final spin formulae (FS0, FS2, FS3, see main text for more details). The two NNLO results are plotted in red (FS0) and purple (FS2), while the default MSA results is plotted in green. Dashed (solid) lines in the 1D (2D) plots indicate 90% credible intervals (regions).

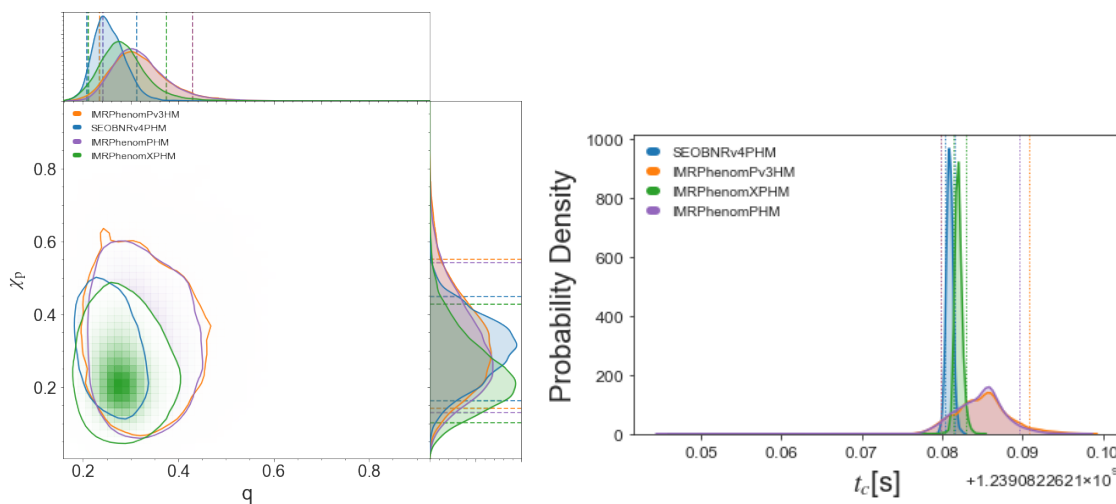


FIG. 17. Left panel: 1D and 2D posterior distributions for the mass ratio and precession spin of the signal recovered with IMRPHENOMPv3HM (orange), SEOBNRv4PHM (blue), IMRPHENOMXPHM in its default version (green) and its “IMRPhenomPHM” version (purple, see main text for details). Dashed (solid) lines indicate 90% credible intervals (regions). Right panel: posterior distributions for the geocentric time of the event, as estimated with the same models shown in the top panel.

Approximant	Run n	Modes (l, $ m $)	PV	FS	PMB	MB	Prior	n_{live}	n_{act}	CPU h	L. eval.	Cost/L. eval. [ms]
IMRPHENOMXAS	1	(2,2)	-	-	-	-	Aligned spin	2048	20	2650	2.23×10^8	42.77
IMRPHENOMXHM	2	D	-	-	-	D	Aligned spin	2048	20	4734	2.90×10^8	58.67
IMRPHENOMXP	3	(2,2)	102	0	-	-	Precessing	2048	5	1042	6.57×10^7	57.13
	4	(2,2)	D	D	-	-	Precessing	2048	5	1273	6.32×10^7	72.45
IMRPHENOMXPHM	5	D	D	D	D	D	Precessing	512	5	1566	1.99×10^7	283.60
	6	D	D	D	D	D	Precessing	1024	5	1531	3.61×10^7	152.55
	7	D	D	D	D	D	Precessing	1024	20	3362	1.31×10^8	92.16
	8	D	D	D	D	D	Uniform m_1 - m_2	1024	30	7082	2.05×10^8	124.17
	9	D	D	D	D	D	Uniform m_1 - m_2	1024	50	8362	3.34×10^8	90.10
	10	D	D	D	D	D	Precessing	1024	60	3696	1.68×10^8	79.33
	11	D	D	D	D	D	Precessing	1500	50	5041	2.31×10^8	78.43
	12	D	D	D	D	D	Precessing	2048	5	1804	7.15×10^7	90.85
	13	D	D	D	10^{-2}	D	Precessing	2048	5	1295	7.19×10^7	64.85
	14	D	D	D	10^{-2}	D	Precessing *	2048	5	1690	7.18×10^7	84.73
	15	D	102	2	D	D	Precessing	2048	5	1336	7.09×10^7	67.83
	16	D	102	0	D	D	Precessing	2048	5	1354	7.10×10^7	68.61
	17	D	223	2	D	D	Precessing	2048	5	1732	7.24×10^7	86.13
	18	No (3,2)	D	D	D	D	Precessing	2048	5	1465	7.21×10^7	73.11
	19	D	D	D	D	D	Volumetric spin	2048	5	2216	7.77×10^7	102.68
	20	D	D	D	D	D	Precessing	2048	10	4275	1.54×10^8	99.44
	21	D	D	D	D	D	Precessing	2048	20	6795	2.66×10^8	91.79
	22	D	D	D	D	D	Uniform m_1 - m_2	2048	30	9070	4.22×10^8	77.25
	23	D	D	D	D	D	Precessing	4096	5	3434	1.46×10^8	84.47
	24	D	D	D	D	D	Precessing	4096	10	6377	2.89×10^8	79.43
	25	D	102	D	D	D	Precessing	4096	10	5132	2.85×10^8	64.75
	26	D	D	D	D	D	Precessing	4096	20	11438	5.71×10^8	72.08
	27	D	D	D	D	D	Uniform m_1 - m_2	4096	30	20187	9.49×10^8	76.57
	28	D	D	D	D	D	Precessing	8192	5	8921	2.95×10^8	108.94
	29	D	D	D	D	D	Precessing, $\chi_1 = 0$	2048	5	1863	6.95×10^7	96.47
	30	D	D	D	D	D	Aligned, $\chi_1 = 0$	2048	5	1726	6.31×10^7	98.40

TABLE VIII. List of the runs performed on GW190412 open data with parallel Bilby *without* distance marginalisation, using different models of the IMRPHENOMX family. For each run, we indicate the LALSuite waveform approximant called along with various waveform settings: the mode content, precessing prescription (PV), final spin version (FS), and multibanding thresholds applied to the evaluation of the Euler angles (PMB) and of the aligned-spin modes (MB); as well as the prior used and the chosen sampler settings (n_{live} and n_{act}). We also provide the computational cost of each run, the number of likelihood evaluations and the mean cost of each evaluation. We use 4kHz GWOSC data for all runs in this table, except for the one marked with an asterisk, which employs 16kHz data.

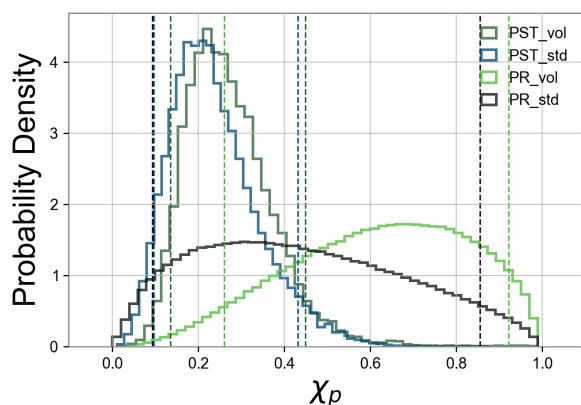


FIG. 18. Induced prior on the effective precessing spin parameter χ_p for two different priors on the component spins: uniform in magnitudes (default prior for this paper) and volumetric (power law with exponent 2 in the spin magnitudes). These correspond to priors “P1” and “P2” in appendix C.1 of [14]. Also shown are the posteriors on χ_p from two IMRP_{HENOMXPHM} runs with the two different prior choices and otherwise identical settings (runs 12 and 19 in Table VIII).

metric prior gives more weight to high χ_p , this result is consistent with the overall picture of no support in the GW190412 data for strongly precessing spins⁶.

In conclusion, the effect of different standard spin prior choices on GW190412 inference is small, with the main conclusions robust against such a change.

We also briefly consider another spin prior as suggested by [17]. They argue that if GW190412 formed from the isolated

binary evolution channel [56, 99], the more massive component BH should have low (or even zero) spin but the secondary BH could have high aligned spin. This suggestion and the effect of spin priors on GW190412 has already been considered in detail by [58]. Here we report on two additional IMRP_{HENOMXPHM} runs with the same settings as above (2048 live points, $n_{\text{act}} = 5$ and without distance marginalisation) but with priors that (i) fix $a_1 = 0$, leaving a_2 and the tilt angles unchanged; (ii) fix $a_1 = 0$ and the secondary spin to be positive aligned (a_2 uniform in [0.0, 0.99], tilt $\theta_2 = 0$). These are runs 29 and 30 in Table VIII.

As suggested by [17] based on reweighting the original LVC posteriors, and first confirmed by [58] from full PE runs, our results also show that it is possible to fit the GW190412 data with zero primary spin, and that the posteriors then prefer a high value for the secondary spin magnitude: $a_2 > 0.52$ or $a_2 > 0.54$ at 90% for the two runs respectively, with both posteriors railing against the upper prior limit. However, this alternative configuration is actually a somewhat worse fit to the data than the standard interpretation of nonzero primary spin⁷: while the run with the same settings and default prior (run no. 12 in Table VIII) reaches network matched filter SNRs of $18.80^{+0.19}_{-0.3}$, for the two modified prior runs these are only $18.52^{+0.23}_{-0.34}$ and $18.53^{+0.19}_{-0.3}$. Correspondingly, we find clear preference against the first and mild preference against the second alternative, with $\log_{10} \mathcal{B}$ of 1.7 ± 0.2 and 1.0 ± 0.2 in favour of the standard interpretation. Hence, our result is consistent with the findings of [58] for other waveforms: also with IMRP_{HENOMXPHM} there appears to be no preference for the scenario of [17] from the GW190412 data, since the reduced prior volume cannot make up for the worse fit to the data of waveforms without primary spin.

-
- [1] R. Abbott *et al.* (LIGO Scientific, Virgo), *Phys. Rev. D* **102**, 043015 (2020), [arXiv:2004.08342 \[astro-ph.HE\]](https://arxiv.org/abs/2004.08342).
- [2] R. Abbott, T. D. Abbott, S. Abraham, F. Acernese, K. Ackley, C. Adams, R. X. Adhikari, V. B. Adya, C. Affeldt, M. Agathos, and *et al.*, *The Astrophysical Journal* **896**, L44 (2020).
- [3] R. Abbott and *et. al.* (LIGO Scientific Collaboration and Virgo Collaboration), *Phys. Rev. Lett.* **125**, 101102 (2020).
- [4] R. Abbott and *et. al.*, *The Astrophysical Journal* **900**, L13 (2020).
- [5] J. Aasi *et al.* (LIGO Scientific), *Class. Quant. Grav.* **32**, 074001 (2015), [arXiv:1411.4547 \[gr-qc\]](https://arxiv.org/abs/1411.4547).
- [6] F. Acernese *et al.* (VIRGO), *Class. Quant. Grav.* **32**, 024001 (2015), [arXiv:1408.3978 \[gr-qc\]](https://arxiv.org/abs/1408.3978).
- [7] G. Pratten, S. Husa, C. García-Quirós, M. Colleoni, A. Ramos-Buades, H. Estellés, and R. Jaume, *Phys. Rev. D* **102**, 064001 (2020).

- [8] C. García-Quirós, M. Colleoni, S. Husa, H. Estellés, G. Pratten, A. Ramos-Buades, M. Mateu-Lucena, and R. Jaume, *Phys. Rev. D* **102**, 064002 (2020).
- [9] C. García-Quirós, S. Husa, M. Mateu-Lucena, and A. Borchers, ArXiv e-prints (2020), [arXiv:2001.10897 \[gr-qc\]](https://arxiv.org/abs/2001.10897).
- [10] G. Pratten *et al.*, ArXiv e-prints (2020), [arXiv:2004.06503 \[gr-qc\]](https://arxiv.org/abs/2004.06503).
- [11] H. Estellés, A. Ramos-Buades, S. Husa, C. García-Quirós, M. Colleoni, L. Haegel, and R. Jaume, “IMRPhenomTP: A phenomenological time domain model for dominant quadrupole gravitational wave signal of coalescing binary black holes,” (2020), [arXiv:2004.08302 \[gr-qc\]](https://arxiv.org/abs/2004.08302).
- [12] H. Estellés and *et. al.*, “IMRPhenomTHM,” (2020), in preparation.
- [13] C. Kalaghatgi, M. Hannam, and V. Raymond, *Phys. Rev. D* **101**, 103004 (2020).
- [14] B. P. Abbott *et al.* (LIGO Scientific, Virgo), *Phys. Rev.* **X9**, 031040 (2019), [arXiv:1811.12907 \[astro-ph.HE\]](https://arxiv.org/abs/1811.12907).
- [15] B. Abbott *et al.* (LIGO Scientific, Virgo), *Astrophys. J. Lett.* **892**, L3 (2020), [arXiv:2001.01761 \[astro-ph.HE\]](https://arxiv.org/abs/2001.01761).
- [16] B. Abbott *et al.* (LIGO Scientific, Virgo), *Astrophys. J.* **882**, L24 (2019), [arXiv:1811.12940 \[astro-ph.HE\]](https://arxiv.org/abs/1811.12940).

⁶ No separate IMRP_{HENOMXPHM} run with modified prior was necessary for this Bayes factor comparison, since for the one-dimensional aligned-spin parameter space a “volumetric” prior is identical to the default uniform prior.

⁷ Note that, for consistency, we will be comparing here only runs performed without distance marginalization.

- [17] I. Mandel and T. Fragos, *Astrophys. J. Lett.* **895**, L28 (2020), [arXiv:2004.09288 \[astro-ph.HE\]](https://arxiv.org/abs/2004.09288).
- [18] U. N. Di Carlo *et al.*, *Monthly Notices of the Royal Astronomical Society* (2020), 10.1093/mnras/staa2286, [arXiv:2004.09525 \[astro-ph.HE\]](https://arxiv.org/abs/2004.09525).
- [19] A. Olejak, M. Fishbach, K. Belczynski, D. Holz, J.-P. Lasota, M. Miller, and T. Bulik, *Astrophys. J.* **901**, L39 (2020), [arXiv:2004.11866 \[astro-ph.HE\]](https://arxiv.org/abs/2004.11866).
- [20] A. S. Hamers and M. Safarzadeh, *Astrophys. J.* **898**, 99 (2020), [arXiv:2005.03045 \[astro-ph.HE\]](https://arxiv.org/abs/2005.03045).
- [21] C. L. Rodriguez *et al.*, *Astrophys. J. Lett.* **896**, L10 (2020), [arXiv:2005.04239 \[astro-ph.HE\]](https://arxiv.org/abs/2005.04239).
- [22] D. Gerosa, S. Vitale, and E. Berti, *Phys. Rev. Lett.* **125**, 101103 (2020), [arXiv:2005.04243 \[astro-ph.HE\]](https://arxiv.org/abs/2005.04243).
- [23] V. De Luca, G. Franciolini, P. Pani, and A. Riotto, *JCAP* **06**, 044 (2020), [arXiv:2005.05641 \[astro-ph.CO\]](https://arxiv.org/abs/2005.05641).
- [24] M. Safarzadeh and K. Hotokezaka, *Astrophys. J. Lett.* **897**, L7 (2020), [arXiv:2005.06519 \[astro-ph.HE\]](https://arxiv.org/abs/2005.06519).
- [25] C. Kimball, C. Talbot, C. P. L. Berry, M. Carney, M. Zevin, E. Thrane, and V. Kalogera, *The Astrophysical Journal* **900**, 177 (2020).
- [26] L. Blanchet, *Living Rev. Rel.* **9**, 4 (2006).
- [27] T. Damour, *Phys. Rev.* **D64**, 124013 (2001).
- [28] T. Damour, A. Nagar, and S. Bernuzzi, *Phys. Rev.* **D87**, 084035 (2013), [arXiv:1212.4357 \[gr-qc\]](https://arxiv.org/abs/1212.4357).
- [29] R. Cotesta, A. Buonanno, A. Bohé, A. Taracchini, I. Hinder, and S. Ossokine, *Phys. Rev.* **D98**, 084028 (2018), [arXiv:1803.10701 \[gr-qc\]](https://arxiv.org/abs/1803.10701).
- [30] E. Berti, V. Cardoso, and C. M. Will, *Physical Review D* **73** (2006), 10.1103/physrevd.73.064030.
- [31] SXS Collaboration, “SXS Gravitational Waveform Database,” <https://www.black-holes.org/waveforms> (2019).
- [32] M. Boyle *et al.*, *Class. Quant. Grav.* **36**, 195006 (2019), [arXiv:1904.04831 \[gr-qc\]](https://arxiv.org/abs/1904.04831).
- [33] S. Khan, F. Ohme, K. Chatziioannou, and M. Hannam, *Phys. Rev.* **D101**, 024056 (2020), [arXiv:1911.06050 \[gr-qc\]](https://arxiv.org/abs/1911.06050).
- [34] S. Ossokine *et al.*, *Phys. Rev. D* **102**, 044055 (2020), [arXiv:2004.09442 \[gr-qc\]](https://arxiv.org/abs/2004.09442).
- [35] S. Husa, S. Khan, M. Hannam, M. Pürrer, F. Ohme, X. Jiménez Forteza, and A. Bohé, *Phys. Rev.* **D93**, 044006 (2016), [arXiv:1508.07250 \[gr-qc\]](https://arxiv.org/abs/1508.07250).
- [36] S. Khan, S. Husa, M. Hannam, F. Ohme, M. Pürrer, X. Jiménez Forteza, and A. Bohé, *Phys. Rev.* **D93**, 044007 (2016), [arXiv:1508.07253 \[gr-qc\]](https://arxiv.org/abs/1508.07253).
- [37] M. Hannam, P. Schmidt, A. Bohé, L. Haegel, S. Husa, *et al.*, *Phys.Rev.Lett.* **113**, 151101 (2014), [arXiv:1308.3271 \[gr-qc\]](https://arxiv.org/abs/1308.3271).
- [38] A. Bohé, M. Hannam, S. Husa, F. Ohme, M. Puerrer, and P. Schmidt, *PhenomPv2 - Technical Notes for LAL Implementation*, Tech. Rep. LIGO-T1500602 (LIGO Project, 2016).
- [39] L. London, S. Khan, E. Fauchon-Jones, C. García, M. Hannam, S. Husa, X. Jiménez-Forteza, C. Kalaghatgi, F. Ohme, and F. Panarale, *Phys. Rev. Lett.* **120**, 161102 (2018), [arXiv:1708.00404 \[gr-qc\]](https://arxiv.org/abs/1708.00404).
- [40] S. Khan, K. Chatziioannou, M. Hannam, and F. Ohme, *Phys. Rev.* **D100**, 024059 (2019), [arXiv:1809.10113 \[gr-qc\]](https://arxiv.org/abs/1809.10113).
- [41] LIGO Scientific Collaboration, “LIGO Algorithm Library - LAL-Suite,” free software (GPL), <https://doi.org/10.7935/GT1W-FZ16> (2020).
- [42] V. Varma, S. E. Field, M. A. Scheel, J. Blackman, D. Gerosa, L. C. Stein, L. E. Kidder, and H. P. Pfeiffer, *Phys. Rev. Research.* **1**, 033015 (2019), [arXiv:1905.09300 \[gr-qc\]](https://arxiv.org/abs/1905.09300).
- [43] T. Islam, S. E. Field, C.-J. Haster, and R. Smith, (2020).
- [44] M. Pürrer, *Classical and Quantum Gravity* **31**, 195010 (2014).
- [45] G. Ashton *et al.*, *Astrophys. J. Suppl.* **241**, 27 (2019), [arXiv:1811.02042 \[astro-ph.IM\]](https://arxiv.org/abs/1811.02042).
- [46] R. Smith, G. Ashton, A. Vajpeyi, and C. Talbot, *arXiv e-prints* (2019), [arXiv:1909.11873 \[gr-qc\]](https://arxiv.org/abs/1909.11873).
- [47] I. Romero-Shaw *et al.*, *Monthly Notices of the Royal Astronomical Society* (2020), 10.1093/mnras/staa2850, [arXiv:2006.00714 \[astro-ph.IM\]](https://arxiv.org/abs/2006.00714).
- [48] J. Veitch *et al.*, *Phys. Rev.* **D91**, 042003 (2015), [arXiv:1409.7215 \[gr-qc\]](https://arxiv.org/abs/1409.7215).
- [49] M. Colleoni, M. Mateu-Lucena, H. Estellés, C. García-Quirós, D. Keitel, G. Pratten, A. Ramos-Buades, and S. Husa, “Data release for paper ‘Towards the routine use of subdominant harmonics in gravitational-wave inference: re-analysis of GW190412 with generation X waveform models’,” Zenodo.org deposit (2020).
- [50] Planck Collaboration, P. A. R. Ade, N. Aghanim, M. Arnaud, M. Ashdown, J. Aumont, C. Baccigalupi, A. J. Banday, R. B. Barreiro, J. G. Bartlett, *et al.*, *A&A* **594**, A13 (2016), [arXiv:1502.01589 \[astro-ph.CO\]](https://arxiv.org/abs/1502.01589).
- [51] P. Ajith, M. Hannam, S. Husa, Y. Chen, B. Brügmann, N. Dorband, D. Müller, F. Ohme, D. Pollney, C. Reisswig, L. Santamaría, and J. Seiler, *Physical Review Letters* **106**, 241101 (2011).
- [52] L. Santamaría, F. Ohme, P. Ajith, B. Bruegmann, N. Dorband, M. Hannam, S. Husa, P. Moesta, D. Pollney, E. L. Reisswig, *et al.*, *Phys. Rev.* **D82**, 064016 (2010).
- [53] P. Schmidt, F. Ohme, and M. Hannam, *Phys. Rev.* **D91**, 024043 (2015), [arXiv:1408.1810 \[gr-qc\]](https://arxiv.org/abs/1408.1810).
- [54] J. Lange, R. O’Shaughnessy, and M. Rizzo, *arXiv e-prints* (2018), [arXiv:1805.10457 \[gr-qc\]](https://arxiv.org/abs/1805.10457).
- [55] D. Wysocki, R. O’Shaughnessy, J. Lange, and Y.-L. L. Fang, *Phys. Rev. D* **99**, 084026 (2019), [arXiv:1902.04934 \[astro-ph.IM\]](https://arxiv.org/abs/1902.04934).
- [56] K. A. Postnov and L. R. Yungelson, *Living Rev. Rel.* **17**, 3 (2014).
- [57] M. J. Benacquista and J. M. B. Downing, *Living Rev. Rel.* **16**, 4 (2013).
- [58] M. Zevin, C. P. Berry, S. Coughlin, K. Chatziioannou, and S. Vitale, *Astrophys. J. Lett.* **899**, L17 (2020), [arXiv:2006.11293 \[astro-ph.HE\]](https://arxiv.org/abs/2006.11293).
- [59] A. Bohé *et al.*, *Phys. Rev.* **D95**, 044028 (2017), [arXiv:1611.03703 \[gr-qc\]](https://arxiv.org/abs/1611.03703).
- [60] R. Cotesta, S. Marsat, and M. Pürrer, *Phys. Rev. D* **101**, 124040 (2020), [arXiv:2003.12079 \[gr-qc\]](https://arxiv.org/abs/2003.12079).
- [61] S. Babak, A. Taracchini, and A. Buonanno, *Phys. Rev.* **D95**, 024010 (2017), [arXiv:1607.05661 \[gr-qc\]](https://arxiv.org/abs/1607.05661).
- [62] Y. Pan, A. Buonanno, A. Taracchini, L. E. Kidder, A. H. Mroué, H. P. Pfeiffer, M. A. Scheel, and B. Szilágyi, *Phys. Rev.* **D89**, 084006 (2014), [arXiv:1307.6232 \[gr-qc\]](https://arxiv.org/abs/1307.6232).
- [63] V. Varma, S. E. Field, M. A. Scheel, J. Blackman, L. E. Kidder, and H. P. Pfeiffer, *Phys. Rev.* **D99**, 064045 (2019), [arXiv:1812.07865 \[gr-qc\]](https://arxiv.org/abs/1812.07865).
- [64] P. Ajith, S. Babak, Y. Chen, M. Hewitson, B. Krishnan, J. T. Whelan, B. Brügmann, P. Diener, J. Gonzalez, M. Hannam, and *et al.*, *Classical and Quantum Gravity* **24**, S689–S699 (2007).
- [65] V. Varma, S. E. Field, M. A. Scheel, J. Blackman, L. E. Kidder, and H. P. Pfeiffer, *Physical Review D* **99** (2019), 10.1103/physrevd.99.064045.
- [66] P. Schmidt, M. Hannam, S. Husa, and P. Ajith, *Phys. Rev.* **D84**, 024046 (2011), [arXiv:1012.2879 \[gr-qc\]](https://arxiv.org/abs/1012.2879).

- [67] P. Schmidt, M. Hannam, and S. Husa, *Phys. Rev.* **D86**, 104063 (2012), arXiv:1207.3088 [gr-qc].
- [68] K. Chatziioannou, A. Klein, N. Yunes, and N. Cornish, *Phys. Rev.* **D95**, 104004 (2017), arXiv:1703.03967 [gr-qc].
- [69] S. Vinciguerra, J. Veitch, and I. Mandel, *Class. Quant. Grav.* **34**, 115006 (2017), arXiv:1703.02062 [gr-qc].
- [70] A. Buonanno, B. R. Iyer, E. Ochsner, Y. Pan, and B. S. Sathyaprakash, *Physical Review D* **80** (2009), 10.1103/physrevd.80.084043.
- [71] T. Damour and A. Nagar, *Physical Review D* **90** (2014), 10.1103/physrevd.90.024054.
- [72] GW Open Science Center, “GW190412 strain data release v2,” <https://doi.org/10.7935/20yv-ka61> (2020).
- [73] M. Vallisneri, J. Kanner, R. Williams, A. Weinstein, and B. Stephens, *Proceedings, 10th International LISA Symposium: Gainesville, Florida, USA, May 18-23, 2014*, *J. Phys. Conf. Ser.* **610**, 012021 (2015), arXiv:1410.4839 [gr-qc].
- [74] R. Abbott *et al.* (LIGO Scientific, Virgo), ArXiv e-prints (2019), arXiv:1912.11716 [gr-qc].
- [75] G. Vajente, Y. Huang, M. Isi, J. C. Driggers, J. S. Kissel, M. J. Szczepanczyk, and S. Vitale, *Phys. Rev. D* **101**, 042003 (2020), arXiv:1911.09083 [gr-qc].
- [76] T. B. Littenberg and N. J. Cornish, *Phys. Rev. D* **91**, 084034 (2015).
- [77] K. Chatziioannou, C.-J. Haster, T. B. Littenberg, W. M. Farr, S. Ghonge, M. Millhouse, J. A. Clark, and N. Cornish, *Phys. Rev. D* **100**, 104004 (2019).
- [78] L. Sun *et al.*, *Classical and Quantum Gravity* (2020), 10.1088/1361-6382/abb14e, arXiv:2005.02531 [astro-ph.IM].
- [79] LIGO Scientific and Virgo Collaborations, “GW190412 posterior samples data release,” <https://dcc.ligo.org/P190412/public> (2020).
- [80] LIGO Scientific Collaboration and Virgo Collaboration, “Gravitational-Wave Candidate Event Database,” (2020).
- [81] R. J. E. Smith, G. Ashton, A. Vajpeyi, and C. Talbot, *Monthly Notices of the Royal Astronomical Society* **498**, 4492 (2020), <https://academic.oup.com/mnras/article-pdf/498/3/4492/33798799/staa2483.pdf>.
- [82] R. Smith, S. E. Field, K. Blackburn, C.-J. Haster, M. Pürrer, V. Raymond, and P. Schmidt, *Phys. Rev. D* **94**, 044031 (2016).
- [83] J. Skilling, in *24th International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering*, AIP Conf. Proc., Vol. 735, edited by R. Fischer, R. Preuss, and U. V. Toussaint (2004) pp. 395–405.
- [84] J. S. Speagle, *Monthly Notices of the Royal Astronomical Society* **493**, 3132 (2020), <https://academic.oup.com/mnras/article-pdf/493/3/3132/32890730/staa278.pdf>.
- [85] C. Hoy and V. Raymond, ArXiv e-prints (2020), arXiv:2006.06639 [astro-ph.IM].
- [86] E. Thrane and C. Talbot, *Publ. Astron. Soc. Austral.* **36**, e010 (2019), arXiv:1809.02293 [astro-ph.IM].
- [87] A. Majtey, P. W. Lamberti, M. T. Martin, and A. Plastino, *The European Physical Journal D - Atomic, Molecular, Optical and Plasma Physics* **32**, 413 (2005).
- [88] S. Kullback and R. A. Leibler, *Ann. Math. Statist.* **22**, 79 (1951).
- [89] F. H. Shaik, J. Lange, S. E. Field, R. O’Shaughnessy, V. Varma, L. E. Kidder, H. P. Pfeiffer, and D. Wysocki, *Phys. Rev. D* **101**, 124054 (2020).
- [90] P. Kumar, J. Blackman, S. E. Field, M. Scheel, C. R. Galley, M. Boyle, L. E. Kidder, H. P. Pfeiffer, B. Szilagy, and S. A. Teukolsky, *Phys. Rev. D* **99**, 124005 (2019).
- [91] S. A. Usman, J. C. Mills, and S. Fairhurst, *The Astrophysical Journal* **877**, 82 (2019).
- [92] B. Abbott *et al.* (KAGRA, LIGO Scientific, VIRGO), *Living Rev. Rel.* **21**, 3 (2018), arXiv:1304.0670 [gr-qc].
- [93] L. Barsotti, P. Fritschel, M. Evans, and S. Gras, *The updated Advanced LIGO design curve*, Tech. Rep. LIGO-T1800044 (LIGO Project, 2018).
- [94] P. Schmidt, I. W. Harry, and H. P. Pfeiffer, ArXiv e-prints (2017), arXiv:1703.01076 [gr-qc].
- [95] I. Harry, S. Privitera, A. Bohé, and A. Buonanno, *Phys. Rev. D* **94**, 024012 (2016).
- [96] A. Ramos-Buades, P. Schmidt, G. Pratten, and S. Husa, *Phys. Rev. D* **101**, 103014 (2020), arXiv:2001.10936 [gr-qc].
- [97] S. Marsat and J. G. Baker, ArXiv e-prints (2018), arXiv:1806.10734 [gr-qc].
- [98] LIGO Scientific Collaboration, Virgo Collaboration, “Gravitational Wave Open Science Center,” <https://www.gwopenscience.org> (2019).
- [99] S. S. Bavera, T. Fragos, Y. Qin, E. Zapartas, C. J. Neijssel, I. Mandel, A. Batta, S. M. Gaebel, C. Kimball, and S. Stevenson, *Astron. Astrophys.* **635**, A97 (2020), arXiv:1906.12257 [astro-ph.HE].